

FIGURE 1
Plasmid sequence of pNC5LSPCEap53 (pMC30B5) for vCP2086

5	1	GCCCTTT	CGTCTCG	CGCGTIT	CGGTGAT	GACGGTG	AAAACCT	CTGACAC	ATGCAGC	TCCCCGA	GACGGTC	
		CGGGAAA	GCAGAGC	GCSCAAA	GCCACTA	CTGCCAC	TTTTGGA	GACTGTG	TACGTGC	AGGGGCT	CTGCCAG	
	71	ACAGCTT	GTCTGTT	AGCGGAT	CGCCGGA	AGCCCGT	CAGGGCG	CGTCAAG	CGGGTGT	CGCGGGT		
		TGTCGAA	CAGACAT	TCGCCTA	CGGCCCT	CGTCTGT	TCGGGCA	GTCGCCG	GCAGTCG	CCCACAA	CCGCCCA	
	141	GTCCGGG	CTGGCTT	AACTATG	CGGCATC	AGAGCAG	ATTGTAT	TGAGAGT	GCACCAT	ATGCGGT	GTGAAAT	
		CAGCCCC	GACCGAA	GCCGTAG	TCTCGTC	TAACATC	ACTCTCA	CTCTGTA	TACGGCA	CACTTTA		
10	211	ACCGCAC	AGATGCG	TAAGGAG	AAAATAC	CGCATCA	GGCGCCA	TTCGCCA	TTCAGGC	TGCGCAA	CTGTTGG	
		TGGCGTG	TCTACGC	ATCCCTC	TTTTATG	GCGTAGT	CCGCGGT	AAGCGGT	AAGTCCG	ACGCGTT	GACAACC	
	281	GAAGGGC	GATCCGC	GTCGGCC	TCTTCGC	TATTACG	CCAGCTG	GCGAAG	GGGGATG	TGCTGCA	AGGCGAT	
		CTTCCCG	CTAGCCA	CGCCCGG	AGAAGCG	ATAATGC	GGTCGAC	CGCTTTC	CCCCTAC	ACGACGT	TCCGCTA	
15	351	TAAGTTG	GGAACCG	CCAGGGT	TTTCCCA	GTCACGA	CGTTGTA	AAACGAG	GGCCAGT	GCCAAGC	TTGGCTG	
		ATTC AAC	CCATTGC	GGTCCCA	AAAGGGT	CAGTGCT	GCAACAT	TTTGCTG	CCGCTGC	CGGTTCC	AACCGAC	
		~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	
												Left Arm
20	421	CAGGTAT	TCTAAAC	TAGGAAT	AGATGAA	ATTATGT	GCAAAGG	AGATACC	TTTAGAT	ATGGATC	TGATTTA	
		GTCCATA	AGATTTG	ATCCTTA	TCTACTT	TAATACA	CGTTTCC	TCTATGG	AAATCTA	TACCTAG	ACTAAAT	
												Left Arm
	491	TTTGTTT	TTTCATA	ATCATAA	TCTAACA	CACATTT	CACTATA	CTATACC	TTCTTGC	ACAAGTC	GCCATTA	
		AAACCAA	AAAGTAT	TAGTATT	AGATTGT	TGTAAAA	GTGATAT	GATATGG	AAGAACG	TGTTTCA	CGGTAAT	
												Left Arm
25	561	GTAGTAT	AGACTTA	TACTTTG	TAACCAT	AGTATAC	TTTAGCG	CGTCATC	TTCTTCA	TCTAAAA	CAGATTT	
		CATCATA	TCTGAAT	ATGAAAC	ATTGGTA	TCATATG	AAATCGC	GCAGTAG	AAGAAGT	AGATTTT	GTCTAAA	
												Left Arm
	631	ACAACAA	TAATCAT	CGTCGTC	ATCTTCA	TCTTCAT	TAAAGTT	TTCATAT	TCAATAA	CTTCTTT	TTCTAAA	
		TGTTGTT	ATTAGTA	GCAGCAG	TAGAAGT	AGAAGTA	ATTTCAA	AAGTATA	AGTTATT	GAAAGAA	AAGATTT	
												Left Arm
30	701	ACATCAT	CTGAATC	AATAAAC	ATAGAAC	GGTATAG	AGCGTTA	ATCTCCA	TTGTAAA	ATATACT	AACGCGT	
		TGTAGTA	GACTTAG	TTATTTG	TATCTTG	CCATATC	TCGCAAT	TAGAGGT	AACATTT	TATATGA	TTGCSCA	
												Left Arm
	771	TGCTCAT	GATGTAC	TTTTTTT	CATTATT	TAGAAAT	TATGCAT	TTTAGAT	CITTATA	AGCGGCC	GTGATTA	
35		ACGAGTA	CTACATG	AAAAAAA	GTAATAA	ATCTTTA	ATACGTA	AAATCTA	GAAATAT	TCGCCGG	CACTAAT	
												~~~~~
												Left Arm
	841	ACTAGTC	ATAAAAA	CCCGGGA	TCGATTG	TAGACTG	GAGATAA	AAACTAT	ATCAGAT	CAACCCC	AACCCAG	
		TGATCAG	TATTTTT	GGGCGCT	AGCTAAG	ATCTGAG	CTCTATT	TTTGATA	TAGTCTC	GTGGGGG	TTGGTGC	
		~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	~~~~~	
40												CEA
												***Ile LeuAla ValGly ValLeuVal-
	911	ACTCAA	TCATGAT	GCCGACA	GTGGCCC	CAGCTGA	GAGACCA	GGAGAAG	TTCCAGA	TGCAGAG	ACTGTGA	
		TGAGGTT	AGTACTA	CGGCTGT	CACCGGG	GTGCACT	CTCTGGT	CCTCTTC	AAGGTCT	ACGTCTC	TGACACT	
												CEA
45		..GlyIle	MetIle	GlyValThr	AlaGly	AlaSer	LeuGlyPro	SerThr	GlySer	AlaSerVal	ThrIle-	
	981	TGCTCTT	GCTATG	GAAATAT	TGCGGCC	AGTAGCC	AAGTATG	AGACAAA	ACAGGCA	TAGGTCT	CGTTATT	
		ACGAGAA	CTGATAC	CTTAATA	ACGCCGG	TCATCGG	TTCAATC	TCTGTTT	TGTCCGT	ATCCAGG	GCAATAA	

1471 GTTGTCTG GAGATGG AGGGCTT GGGCAGC TCCGCGG AAACAGT TATTGTT TTAACGT TAGTCCT GCTGTGA  
CAACGAC CTCTACC TCCCGAA CCCGTCG AGGCGCC TTTGTCA ATAACAA AATTGAC ATCAGGA CGACACT  
CEA

5 1541 AsnSerSer IleSer ProLys ProLeuGlu AlaSer ValThr IleThrLys ValThr ThrArg SerHisGly-  
CCACTGG CTGAGTT ATTGGCC TGGCAAG TATAGAG TCCGCTG TTCTTCT CAGTTAT GTTGTCT ATAATA  
GGTGACC GACTCAA TAACCG ACCGTTT ATATCTC AGGCGAC AAGAAGA GTCAATA CAACGAA TATTTAT  
CEA

10 1611 ..SerAla SerAsn AsnAlaGln CysThr TyrLeu GlySerAsn LysGlu ThrIle AsnSerIle PheLeu-  
ACTCTTG AGTATGC TGCTGAA TGTTTCC ATCAATC AGCCAGG AGTACTG TGCAGGG GGGTTGG ATGCTGC  
TGAGAAC TCATACG ACGACTT ACAAAGG TAGTTAG TCGGTCC TCATGAC ACGTCCC CCAACCC TACGACG  
CEA

15 1681 .GluGln ThrHisGln GlnIle AsnGly AspIleLeu TrpSer TyrGln AlaProPro AsnSer AlaAla  
ATGGCAA GAAAGGC TCAAGTT CACGCCG GGACGGT AGTAGGT GTATGAT GGAGATA TAGTTGG GTCGTCT  
TACCGTT CTTTCGG AGTTCAA GTGCGGC CCTGCCA TCATCCA CATACTA CCTCTAT ATCAACC CAGCAGA  
CEA

20 1751 HisCysSer LeuSer LeuAsn ValGlyPro ArgTyr TyrThr TyrSerPro SerIle ThrPro AspAspPro-  
GGGCCAT ACAAAC ATTAAGG ATAACAG GGTCCGA GTGATCA ACGGATA ATTCATT CTGAATG CCACACT  
CCCGTA TGTTTG TAATTCC TATGTG CAGCCT CACTAGT TGCCCTAT TAAGTAA GACTTAC GGTGTGA  
CEA

25 1821 ..GlyTyr LeuVal AsnLeuIle ValPro AspSer HisAspVal SerLeu GluAsn GlnIleGly CysGlu-  
CATAAGG TCCTACA TCATTGC GAGTAAC GGACAGG AGTGTCA ATGTGCG GTTATCA TTAGACA ACTGCAA  
GTATTCC AGGATGT AGTAACG CTCATTG CTTGTCC TCACAGT TACACGC CAATAGT AATCTGT TGACGTT  
CEA

30 1891 .TyrPro GlyValAsp AsnArg ThrVal SerLeuLeu ThrLeu ThrArg AsnAspAsn SerLeu GlnLeu  
GCGTGGG CTAACCG GCAAACT TTGGTTA TTGACCC ACCATAA ATAAGTG GTATTTT GAATCTC TGGTCTC  
CGCACCC GATTGGC CGTTTGA AACCAAT AACTGGG TGGTATT TATTCAC CATAAAA CTTAGAG ACCGAGT  
CEA

35 1961 ArgProSer ValPro LeuSer GlnAsnAsn ValTrp TrpLeu TyrThrThr AsnGln IleGlu ProGluCys-  
CAAGTTA ATGCAAC TGCGTCC TCATCCT CAACTGG GTTAGAA TTGTAC TAGTTAT GAATGGT TTTGGTG  
GTTCAAT TACGTTG ACGCAGG AGTAGGA GTTGACC CAATCTT AACAATG ATCAATA CTTACCA AAACCAC  
CEA

40 2031 ..ThrLeu AlaVal AlaAspGlu AspGlu ValPro AsnSerAsn AsnSer ThrIle PheProLys ProPro-  
GCTCATA CACGGTA ATCGTCG TCACGGT TGTGCGG TTGAGTC CGGTGTC GCTATTG TGAGCTT GGCACGT  
CGAGTAT GTGCCAT TAGCAGC AGTGCCA ACACGCC AACTCAG GCCACAG CGATAAC ACTCGAA CCGTGCA  
CEA

45 2101 .GluTyr ValThrIle ThrThr ValThr ThrArgAsn LeuGly ThrAsp SerAsnHis AlaGln CysThr  
GTAGCAT CCACTAT TGTCAC GGTAAATA TTGGGAA TGAACAG TTCTTGG GTGGACT GTTGAA GGTGCA  
CATCCTA GGTGATA ACAAGTG CCATTAT AACCTTT ACTTGTC AAGGACC CACCTGA CAACCTT TCACGGT  
CEA

50 2171 TyrSerGly SerAsn AsnVal ThrIleAsn ProIle PheLeu GluGlnThr SerGln GlnPhe ThrGlyAsn-  
TTGACAA ACCAGCT GTATTGG GCGGGAG GATTGCT AGCGGCA TGACAGC TCAGATT CAGATTT TCCCTG  
AACTGTT TGGTGCA CATAACC CGCCCTC CTAACGA TCGCCGT ACTGTG AGTCTAA GTCTAAA AGGGGAC  
CEA

55 2241 ..ValPhe TrpSer TyrGlnAla ProPro AsnSer AlaAlaHis CysSer LeuAsn LeuAsnGlu GlySer-  
ATCTATA GCTTGTG TTTAGAG GGCTGAT TGTAGGA GCATCGG GTCCGTA AAGCAGC TTGAGAA TCACTGA  
TAGATAT CGAACAC AATCTC CCGACTA ACATCCT CGTAGCC CAGGCAT TTCGTGC AACTCTT GTGACT  
CEA

60 2311 .ArgTyr SerThrAsn LeuPro SerIle ThrProAla AspPro GlyTyr LeuValAsn LeuIle ValSer  
ATCAGAC CTCCTGG CGCTGAC TGGATT TGGGTTT CGCATTT GTAGCTT GCTGTGT CGTCTCT GGTGACG  
TAGTCTG GAGGACC GCGACTG ACCTAAA ACCCAAA GCGTAAA CATCGAA CGACACA GCAAGGA CAGTGC  
CEA

65 2381 AspSerArg ArgAla SerVal ProAsnGln ThrGlu CysLys TyrSerAla ThrAsp AsnArg ThrValAsn-  
TTAAACA GGGTCAG AGTTCTA TTTCCGT TGCTGAG TTGGAGT CTAGGGG ACACAGG CAGGGAC TGGTGT  
AATTTGT CCCAGTC TCAAGAT AAAGGCA ACGACTC AACCTCA GATCCCC TGTGTCC GTCCCTG ACCAACA  
CEA

70 2451 ..PheLeu ThrLeu ThrArgAsn GlyAsn SerLeu GlnLeuArg ProSer ValPro LeuSerGln AsnAsn-  
TCACCCA CCAGAGA TATGTTG CGTCTTG AGTTTCG GGCTCGC ATGTAAA AGCGACG GCATCTT TGTCTTC  
AGTGGGT GGTCTCT ATACAAC GCAGAAC CCGAGCG TACATTT TCGCTGC CGTAGAA ACAGGAG  
CEA

75 2521 .ValTrp TrpLeuTyr ThrAla AspGln ThrGluPro GluCys ThrPhe AlaValAla AspLys AspGlu  
GACAGGC TTACTAT TATTGGA GCTAATA GAAGGCT TAGGGAG TTCCGGG TATACCC GGAACCT GCCAGTT  
CTGTCCG AATGATA ATAACCT CGATTAT CTTCCGA ATCCCTC AAGGCC ATATGGG CCTTGAC CGGTCAA  
CEA

80 2591 ValProLys SerAsn AsnSer SerIleSer ProLys ProLeu GluProTyr ValArg PheGln GlyThrAla-  
GCTTCTT CATTAC AAGATCT GACTTTA TGACGTG TAGGGTG TAGAATC CTGTGTC ATTCTGG ATGATGT  
CGAAGAA GTAAGTG TTCTAGA CTGAAAT ACTGCAC ATCCAC ATCTTAG GACACAG TAAGACC TACTACA  
CEA

85 2661 ..GluGlu AsnVal LeuAspSer LysIle ValHis LeuThrTyr PheGly ThrAsp AsnGlnIle IleAsn-  
TCTGGAT CAGCAGG GATGCAT TGGGGTA TATTATC TCTCGAC CACTGTA TGCGGGC CCTGGGG TAGCTTG  
AGACCTA GTCGTCC CTACGTA ACCCAT ATAATAG AGAGCTG GTGACAT ACGCCCG GGACCCC ATCGAAC  
CEA

90 2731 .GlnIle LeuLeuSer AlaAsn ProTyr IleIleGlu ArgGly SerTyr AlaProGly ProThr AlaGln  
TTGAGTT CCTATTA CATATCC TATAATT TGACGGT TGCCATC CACTCTT TCACCTT TGTACCA GCTGTAG  
AACTCAA GGATAAT GTATAGG ATATTAA ACTGCCA ACGGTAG GTGAGAA AGTGGAA ACATGGT CGACATC

CEA  
 2801 GlnThrGly IleVal TyrGly IleIleGln ArgAsn GlyAsp ValArgGlu GlyLys TyrTrp SerTyrGly·  
 CCAAAAA GATGCTG GGGCAGA TTGTGGA CAAGTAG AAGCACC TCCTTCC CCTCTGC GACATTG AACGGCG  
 5 GGTTTT CTACGAC CCCGTCT AACACCT GTTCATC TTCGTGG AGGAAGG GGAGACG CTGTAAC TTGCCGC

CEA  
 2871 ..PheLeu HisGln ProLeuAsn HisVal LeuLeu LeuValGlu LysGly GluAla ValAsnPhe ProThr·  
 TGGATTG AATAGTG AGCTTGG CAGTGGT GGGCGGG TTCCAGA AGGTTAG AAGTGAG GCTGTGA GCAGGAG  
 ACCTAAG TTATCAC TCGAACC GTCACCA CCCGCCC AAGGTCT TCCAATC TTCATC CGACACT CGTCTC

CEA  
 10 2941 .SerGlu IleThrLeu LysAla ThrThr ProProAsn TrpPhe ThrLeu LeuSerAla ThrLeu LeuLeu  
 CCTCTGC CAGGGGA TGCACCA TCTGTGG GGAGGGG CCGAGGG AGACTCC ATTATTT ATATTCC AAAAAA  
 GGAGACG GTCCCCT ACGTGGT AGACACC CCTCCCC GGCTCCC TCTGAGG TAATAAA TATAAGG TTTTATT  
 ~~~~~~  
 E/L Promoter
 ~~~~~~

CEA  
 15 ArgGlnTrp ProIle CysTrp ArgHisPro ProAla SerPro SerGluMet  
 ~~~~~~  
 H6 promoter
 ~~~~~~

20 3011 AAAAAA AAATTTT AATTTT GTCGACC TGCAGCT CGACGGA TCCCCC GGGTCTT TTATTCT ATACTTA  
 TTTTAT TTAAAG TTAAAA CAGCTGG ACGTCGA GCTGCCT AGGGGG CCCAGA AATAAGA TATGAAT  
 ~~~~~~  
 E/L Promoter
 ~~~~~~  
 H6 promoter  
 ~~~~~~

25 3081 AAAAGTG AAAATAA ATACAAA GGTTCTT GAGGGTT GTGTAA ATTGAAA GCGAGAA ATAATCA TAAATTA
 TTTTAC TTTTATT TATGTT CCAAGAA CTCCCAA CACAAT TAACCTT CGCTCTT TATTAGT ATTTAAT
 ~~~~~~  
 p53  
 ~~~~~~

30 H6 promoter
 ~~~~~~

3151 MetGlu GluProGln SerAsp ProSer ValGluPro·  
 TTTCATT ATCGCGA TATCCGT TAAGTTT GTATCGT AATGGAG GAGCCGC AGTCAGA TCCTAGC GTCGAGC  
 35 AAAGTAA TAGCGCT ATAGGCA ATTCAAA CATAGCA TTACCTC CTCGGCG TCAGTCT AGGATCG CAGCTCG  
 ~~~~~~  
 p53
 ~~~~~~

3221 ..ProLeu SerGln GluThrPhe SerAsp LeuTrp LysLeuLeu ProGlu AsnAsn ValLeuSer ProLeu·  
 CCCCTCT GAGTCAG GAAACAT TTTTACA CCTATGG AACTAC TTCCTGA AAACAAC GTTCTGT CCCCTT  
 40 GGGGAGA CTCAGTC CTTTGTAA AAAGTCT GGATACC TTGTATG AAGGACT TTTGTG CAAGACA GGGGGAA  
 ~~~~~~  
 p53
 ~~~~~~

3291 .ProSer GlnAlaMet AspAsp LeuMet LeuSerPro AspAsp IleGlu GlnTrpPhe ThrGlu AspPro  
 GCGTCC CAAGCAA TGGATGA TTGATG CTGTCCC CGGACGA TATGAA CAATGGT TCACTGA AGACCCA  
 45 CCGCAGG GTTCGTT ACCTACT AAACCTAC GACAGGG GCCTGCT ATAACCT GTTACCA AGTGACT TCTGGGT  
 ~~~~~~  
 p53
 ~~~~~~

3361 GlyProAsp GluAla ProArg MetProGlu AlaAla ProPro ValAlaPro AlaPro AlaAla ProThrPro·  
 GGTCCAG ATGAAGC TCCCAGA ATGCCAG AGGCTGC TCCCCC GTGGCCC CTGCACC AGCAGCT CCTACAC  
 50 CCAGGTC TACTTCG AGGGTCT TACGGTC TCCGACG AGGGGGG CACCGGG GACGTGG TCGTCGA GGATGTG  
 ~~~~~~  
 p53
 ~~~~~~

3431 ..AlaAla ProAla ProAlaPro SerTrp ProLeu SerSerSer ValPro SerGln LysThrTyr GlnGly·  
 CGGCGCG CCCTGCA CCAGCCC CCTCTG GCCCTG TCATCTT CTGTCCC TTCCAG AAAACCT ACCAGGG  
 55 GCGCCG GGGACGT GTTCGGG GGAGGAC CGGGGAC AGTAGAA GACAGGG AAGGGTC TTTTGA TGGTCCC  
 ~~~~~~  
 p53
 ~~~~~~

3501 .SerTyr GlyPheArg LeuGly PheLeu HisSerGly ThrAla LysSer ValThrCys ThrTyr SerPro  
 CAGCTAC GGTTTCC GTCTGGG CTTCTTG CATCTG GGACAGC CAAGTCT GTGACTT GCACGTA CTCCCT  
 60 GTCGATG CCAAAGG CAGACCC GAAGAAC GTAAGAC CCGTCTG GTTCTGA CACTGAA CGTGCAT GAGGGGA  
 ~~~~~~  
 p53
 ~~~~~~

3571 AlaLeuAsn LysMet PheCys GlnLeuAla LysThr CysPro ValGlnLeu TrpVal AspSer ThrProPro·  
 GCCCTCA ACAAGAT GTTTTGC CAACTGG CCAAGAC CTGCCCT GTGCAGC TGTGGGT TGATTCC ACACCCC  
 65 CGGAGT TGTCTA CAAAACG GTTGACC GGTTCTG GACGGGA CACGTCT ACACCCA ACTAAGG TGTGGGG  
 ~~~~~~  
 p53
 ~~~~~~

3641 ..ProGly ThrArg ValArgAla MetAla IleTyr LysGlnSer GlnHis MetThr GluValVal ArgArg·  
 CGCCCG CACCCG GTCCGCG CCATGGC CATCTAC AAGCAGT CACAGCA CATGACG GAGGTG TGAGGGG  
 70 GCGGGC GTGGGCG CAGGCGC GGTACCG GTAGATG TTCGTCA GTGTCGT GTACTGC CTCCAAC ACTCCGC  
 ~~~~~~  
 p53
 ~~~~~~

3711 .CysPro HisHisGlu ArgCys SerAsp SerAspGly LeuAla ProPro GlnHisLeu IleArg ValGlu  
 CTGCCCG CACCATG AGCGCTG CTCAGAT AGCGATG GTCTGGC CCTCCT CAGCATC TTATCCG AGTGGAA  
 GAGGGG GTGGTAC TCGCGAC GAGTCTA TCGCTAC CAGACCG GGGAGGA GTCGTAG AATAGGC TCACCTT

p53

~~~~~

5 3781 GlyAsnLeu ArgVal GluTyr LeuAspAsp ArgAsn ThrPhe ArgHisSer ValVal ValPro TyrGluPro·
GGAAATT TCGTGT GGAGTAT TTGGATG ACAGAAA CACTTTT CGACATA GTGTGGT GGTGCC TATGAGC
CCTTTAA ACGCACA CCTCATA AACCTAC TGTCTTT GTGAAA GCTGTAT CACACCA CCACGGG ATACTCG

p53

10 3851 ..ProGlu ValGly SerAspCys ThrThr IleHis TyrAsnTyr MetCys AsnSer SerCysMet GlyGly·
CGCCTGA GGTGGC TCTGACT GTACCAC CATCCAC TACAACT ACATGTG TAACAGT TCCTGCA TGGGCGG
GCGGACT CCAACCG AGACTGA CATGGTG GTAGGTG ATGTTGA TGTACAC ATTGTCA AGGACGT ACCCGCC

p53

15 3921 .MetAsn ArgArgPro IleLeu ThrIle IleThrLeu GluAsp SerSer GlyAsnLeu LeuGly ArgAsn
CATGAAC CGGAGGC CCATCCT CACCAC ATCACAC TGGAAGA CTCCAGT GGTAAATC TACTGGG ACGGAAC
GTAATTG GCCTCCG GGTAGGA GTGGTAG TAGTGTG ACCTTCT GAGGTCA CCATTAG ATGACCC TGCCTTG

p53

20 3991 SerPheGlu ValArg ValCys AlaCysPro GlyArg AspArg ArgThrGlu GluGlu AsnLeu ArgLysLys·
AGCTTTG AGGTGCG TGTTTGT GCCTGTC CTGGGAG AGACCGG CGCACAG AGGAAGA GAATCTC CGCAAGA
TCGAAAC TCCACGC ACAACA CGGACAG GACCTC TCTGGCC GCGTGTC TCCTTCT CTTAGAG GCGTTCT

p53

25 4061 ..GlyGlu ProHis HisGluLeu ProPro GlySer ThrLysArg AlaLeu ProAsn AsnThrSer SerSer·
AAGGGGA GCCTCAC CACGAGC TGCCCCC AGGGAGC ACTAAGC GAGCACT GCCAAC AACACCA GCTCCTC
TTCCCCC CGGAGTG GTGCTCG ACGGGGG TCCCTCG TGATTG CTCGTGA CGGGTTG TTGTGGT CGAGGAG

p53

30 4131 .ProGln ProLysLys LysPro LeuAsp GlyGluTyr PheThr LeuGln IleArgGly ArgGlu ArgPhe
TCCCCAG CCAAAGA AGAAACC ACTGGAT GGAGAAT ATTTTAC CCTTCAG ATCCGTG GCGGTGA GCGCTTC
AGGGGTC GGTCTCT TCTTTG TGACCTA CTTCTTA TAAAGTG GGAAGTC TAGGCAC CCGCACT CGCGAAG

p53

35 4201 GluMetPhe ArgGlu LeuAsn GluAlaLeu GluLeu LysAsp AlaGlnAla GlyLys GluPro GlyGlySer·
GAGATGT TCCGAGA GCTGAAT GAGGCTT TGGAACT CAAGGAT GCCCAGG CTGGGAA GGAGCCA GGGGGGA
CTCTACA AGGCTCT CGACTTA CTCCGGA ACCTTGA GTTCCTA CGGGTCC GACCTTT CCTCGGT CCCCCCT

p53

40 4271 ..ArgAla HisSer SerHisLeu LysSer LysLys GlyGlnSer ThrSer ArgHis LysLysLeu MetPhe·
GCAGGGC TCACTCC AGCCACC TGAAATC CAAAAG GGTCACT CTACCTC CCGCCAT AAAAAAC TCATGTT
CGTCCC AGTGAGG TCGTGG ACTTCAG GTTTTTC CAGTCA GATGGAG GCGGTA TTTTGTG AGTACAA

p53

45 4341 .LysThr GluGlyPro AspSer Asp\*\*\*
CAAGACA GAAGGGC CTGACTC AGACTGA ACGCGTT TTTTATC CCGGGCT CGAGGGT ACCGGAT CCTTTTT
GTTCTGT CTTCCCG GACTGAG TCTGACT TGCGCAA AAAATAG GGCCCGA GCTCCCA TGGCCTA GGAAAAA
4411 ATAGCTA ATTAGTC ACGTACC TTTGAGA GTACCAC TTCAGCT ACCTCTT TTGTGTC TCAGAGT AACCTTC
TATCGAT TAATCAG TGCATGG AAACCTCT CATGGTG AAGTCGA TGGAGAA AACACAG AGTCTCA TTGAAG

~~~~~

50 4481 TTTAATC AATTCCA AAACAGT ATATGAT TTTCCAT TTCTTTC AAAGATG TAGTTTA CATCTGC TCCTTTG  
AAATTAG TTAAGGT TTTGTCA TATACTA AAAGGTA AAGAAAG TTTCTAC ATCAAAT GTAGACG AGGAAAC

Right Arm

55 4551 TTGAAAA GTAGCCT GAGCACT TCTTTTC TACCATG AATTACA GCTGGCA AGATCAA TTTTTC CAGTTCT  
AACTTTT CATCGGA CTCGTGA AGAAAAG ATGGTAC TTAATGT CGACCGT TCTAGTT AAAAAAG GTCAGA

Right Arm

60 4621 GGACATT TTATTTT TTTTAAG TAGTGTG CTACATA TTTCAAT ATTTCCA GATTGTA CAGCGAT CATTAAA  
CCTGTAA AATAAAA AAAATTC ATCACAC GATGTAT AAAGTTA TAAAGGT CTAACAT GTCGCTA GTAATTT

Right Arm

65 4691 GGAGTAC GTCCCAT GTTATCC AGCAAGT CAGTATC AGCACCT TTGTTC ATAGAAG TTTAACC ATTGTTA  
CCTCATG CAGGGTA CAATAGG TCGTTCA GTCATAG TCGTGG AACAAGT TATCTTC AAATTGG TAACAAT

Right Arm

70 4761 AATTTTT ATTTGAT ACGGCTA TATGTAG AGGAGTT AACCGAT CCGTGTG TGAAATA TCTACAT CCGCCGA  
TTAAAAA TAAACTA TGCCGAT ATACATC TCCTCAA TTGGCTA GGCACAA ACTTTAT AGATGTA GCGGCT

Right Arm

4831 ATGAGCC AATAGAA GTTTAAC CAAATTA ACTTTGT TAAGGTA AGCTGCC AAACACA AAGGAGT AAAGCCT  
TACTCGG TTATCTT CAAATG GTTTAAT TGAAACA ATTCAT TCGACGG TTTGTGT TTCCTCA TTCGGA

Right Arm

4901 CCGCTGT AAAGAAC ATTGTTT ACATAGT TATTCTT CAACAGA TCTTTC CTATTTT GTAGTCG TCTCTCA  
GGCGACA TTTCTTG TAACAAA TGTATCA ATAAGAA GTTGTCT AGAAAGT GATAAAA CATCAGC AGAGAGT

Right Arm

4971 ACACCGC ATCATGC AGACAAG AAGTTGT GCATTCA GTAACCT CAGGTTT AGCTCCA TACCTCA TCAAGAT  
TGTGGCG TAGTACG TCTGTTC TTCAACA CTAAGT CATTGAT GTCCAAA TCGAGGT ATGGAGT AGTTCTA

Right Arm

5041 TTTTATA GCCTCGG TATTCTT GAACATT ACAGCCA TTTCAAG AGGAGAT TGTAGAG TACCATA TTCCGTG

		AAAAATAT	CGGAGCC	ATAAGAA	CTTGTA	TGTCGGT	AAAGTTC	TCCTCTA	ACATCTC	ATGGTAT	AAGGCAC
						Right Arm					
5	5111	TTAGGGT	CGAATCC	ATTGTCC	AAAAACC	TATTTAG	AGATGCA	TTGTCA	TATCCAT	GATAGCC	TCACAGA
		AATCCCA	GCITAGG	TACACGG	TTTTTGG	ATAAATC	TCTACGT	AACAGTA	ATAGGTA	CTATCGG	AGTGTCT
						Right Arm					
	5181	CGTATAT	GTAAGCC	ATCTTGA	ATGTATA	ATTTTGT	TGTTTTT	AACAACC	GCTCGTG	AACAGCT	TCTATAC
		GCATATA	CATTTCG	TAGAACT	TACATAT	TAAAACA	ACAAAAG	TTGTTTG	CGAGCAC	TTGTCSA	AGATATG
						Right Arm					
10	5251	TTTTTCA	TTTTCTT	CATGATT	AATATAG	TTTACGG	AATATAA	GTATACA	AAAAGTT	TATAGTA	ATCTCAT
		AAAAAGT	AAAAGAA	GTACTAA	TTATATC	AAATGCC	TTATATT	CATATGT	TTTTCAA	ATATCAT	TAGAGTA
						Right Arm					
	5321	AATATCT	GAAACAC	ATACATA	AAACATG	GAAAGAT	TACACGA	TGTCGTT	GAGATAA	ATGGCTT	TTTATTT
		TTATAGA	CTTTGTG	TATGTAT	TTTGTAC	CTTCTTA	ATGTGCT	ACAGCAA	CTCTATT	TACCGAA	AAATAAC
						Right Arm					
15	5391	TCATAGT	TTACAAA	TTCCGAG	TAATCTT	CATCTTT	TACGAAT	ATTCGAG	AATCTGT	TTTATCC	AACCACT
		AGTATCA	AATGTTT	AAGCGTC	ATTAGAA	TAGAAA	ATGCTTA	TAACGTC	TTAGACA	AAATAGG	TTGTCTA
						Right Arm					
	5461	GATTTTT	GTATAAT	ATAACTG	GTATCCT	ATCTTCC	GATAGAA	TGCTGTT	ATTTAAC	ATTTTTG	CACCTAT
		CTAAAAA	CATATTA	TATTGAC	CATAGGA	TAGAAGG	CTATCTT	ACGACAA	TAAATTG	TAAAAAC	GTGGATA
						Right Arm					
20	5531	TAAGTTA	CATCTGT	CAATACC	ATCTTTC	CAACTGA	CTTTATG	TAACGAT	GCGAAAT	AGCATTT	ATCACTA
		ATTCAAT	GTAGACA	GTTTAGG	TAGAAAG	GTTGACT	GAAATAC	ATTGCTA	CGCTTTA	TCCGAAA	TAGTGAT
						Right Arm					
25	5601	TGTCGTA	CCCAATT	ATCATGA	CAAGATT	CTCTTAA	ATACGTA	ATCTTAT	TATCTCT	TGCATAT	TCGTAAT
		ACAGCAT	GGGTTAA	TAGTACT	GTTCTAA	GAGAAAT	TATGCAT	TAGAATA	ATAGAGA	ACGTATA	AGCATTA
						Right Arm					
	5671	AGTAATT	GTAAAGA	GTATACG	ATAACAG	TATAGAT	ATACACG	TGATATA	AATATTT	AACCCCA	TTCTCTA
		TCATTAA	CATTCTT	CATATGC	TATTGTC	ATATCTA	TATGTGC	ACTATAT	TTATAAA	TTGGGGT	AAGCACT
						Right Arm					
30	5741	GTAAAA	AATTACG	ATATTAC	ATTTCTT	TTTATTA	TTTTTTT	GTTTTAG	TTATTTG	TTAGGTT	ATACAAA
		CATTTTA	TTAATGC	TATAATG	TAAAGGA	AAATAAT	AAAAATA	CAAAATC	AAATAAC	AAATCAA	TATGTTT
						Right Arm					
	5811	AATTATG	TTTATTT	GTGTATA	TTTAAAG	CGTCGTT	AAGAATA	AGCTTAG	TTAACAT	ATATATC	CTTAGGT
		TTAATAC	AAATAAA	CACATAT	AAATTTT	GCAGCAA	TTCTTAT	TCGAATC	AATTGTA	TAAATAG	GAATCCA
						Right Arm					
35	5881	TTGTAG	TATTTGA	ATCCTTT	CTTTAAA	TGGATTA	TTTTTCC	AATGCAT	ATTTATA	GCTTCAT	CCAAAGT
		AAACATC	ATAAACT	TAGGAAA	GAAATTT	ACCTAAT	AAAAAGG	TTACGTA	TAAATAT	CGAAGG	GGTTTCA
						Right Arm					
40	5951	ATAACAT	TTAACAT	TCAGAA	TGCGCGG	GCAATTC	AATTCGT	AATCATG	GTCTATG	CTGTTTC	CTGTGTG
		TATTGTA	AATTGTA	AGTCTTA	ACGCGCG	CTTAAAG	TTAAGCA	TTAGTAC	CAGTATC	GACAAAG	GACACAC
						Right Arm					
						Right Arm					
45	6021	AAATTGT	TATCCGC	TCACAAT	TCCACAC	AACATAC	GAGCCGG	AAGCATA	AAGTGTA	AAGCCTG	GGGTGCC
		TTTAAAC	ATAGGCG	AGTGTTA	AGGTGTG	TTGTATG	CTCGGCC	TTCGTAT	TTCCACAT	TTCCGAC	CCCACGG
	6091	TAAATAG	TGAGCTA	ACTCACA	TTAATTG	CGTTGCG	CTCACTG	CCCGCTT	TCCAGTC	GGGAAAC	CTGTCTG
		ATTACTC	ACTCGAT	TGAGTGT	AATTAA	GCAACGC	GAGTGAC	GGGCGAA	AGGTCAG	CCCTTTG	GACAGCA
	6161	GCCAGCT	GCATTAA	TGAATCG	GCCAACG	CGCGGGG	AGAGGCG	GTTTGCG	TATTGGG	CGCTCTT	CCGCTTC
		CGGTCSA	CGTAATT	ACTTAGC	CGGTTCG	GCGCCCC	TCTCCGC	CAAACGC	ATAACCC	GCGAGAA	GGCGAAG
	6231	CTCGCTC	ACTGACT	CGCTGCG	CTCGGTC	GTTCCGC	TGCGGCG	AGCGGTA	TCAGCTC	ACTCAAA	GGCGGTA
		GAGCGAG	TGACTGA	GCGACGC	GAGCCAG	CAAGCCG	ACGCCGC	TCCGCTT	AGTCGAG	TGAGTTT	CCGCCAT
	6301	ATACGGT	TATCCAC	AGAATCA	GGGGATA	ACGCAGG	AAAGAAC	ATGTGAG	CAAAAGG	CCAGCAA	AAGGCCA
		TATGCCA	ATAGGTG	TCTTAGT	CCCCAT	TGCGTCC	TTCTTTG	TACACTC	GTTTCTC	TTCCGGT	TTCCGGT
	6371	GGAACCG	TAAAAAG	GCCGCGT	TGCTGGC	GTTTTTC	CATAGGC	TCCGCC	CCCTGAC	GAGCATC	ACAAAAA
		CCTTGCG	ATTTTTT	CGGCGCA	ACGACCG	CAAAAAG	GTATCCG	AGGCGGG	GGGACTG	CTCGTAG	TGTTTTT
	6441	TCGACGC	TCAAGTC	AGAGGTG	GCGAAAC	CCGACAG	GACTATA	AAGATAC	CAGGCGT	TTCCCCC	TGGAAGC
		AGCTGCG	AGTTCAG	TCTCCAC	CGCTTTG	GGCTGTC	CTGATAT	TTCTATG	GTCCGCA	AAGGGGG	ACCTTCG
	6511	TCCTTCG	TGCGCTC	TCCTGTT	CCGACCC	TGCCGCT	TACCGGA	TACCTGT	CCGCCTT	TCTCCCT	TCGGGAA
		AGGGAGC	ACGCGAG	AGGACAA	GGCTGGG	ACGGCGA	ATGGCCT	ATGGACA	GGCGGAA	AGAGGGA	AGCCCTT
	6581	GCGTGCG	GCCTTCT	CATAGCT	CACGCTG	TAGGTAT	CTCAGTT	CGTGTA	GGTCGTT	CGCTCCA	AGCTGGG
		CGCACCG	CGAAAGA	GTATCGA	GTGCGAG	ATCCATA	GAGTCAA	GCCACAT	CCAGCAA	GCGAGGT	TCGACCC
	6651	CTGTGTG	CACGAAC	CCCCCGT	TCAGCCC	GACCCGT	ATCCGGT	AATCCGT	AACTATC	GTCTTGA	GTCCAAC
		GACACAC	GTGCTTG	GGGGGCA	AGTCGGG	CTGGCGA	CGCGGAA	TAGGCCA	TTGATAG	CAGAACT	CAGGTTG
	6721	CCGGTAA	GACACGA	CTTATCG	CCACTGG	CAGCAGC	CACTGGT	AACAGGA	TTAGCAG	AGCGAGG	TATGTAG
		GGCCATT	CTGTGCT	GAATAGC	GGTGACC	GTCGTCG	GTGACCA	TTGTCC	AATCGTC	TCGCTCC	ATACATC
	6791	GCGGTGC	TACAGAG	TTCTTGA	AGTGGTG	GCCTAAC	TACGGCT	ACACTAG	AAGGACA	GTATTTG	GTATCTG
		CGCCACG	ATGTCTC	AAGAACT	TCACCAC	CGGATTG	ATGCCGA	TGTGATC	TTCTGTT	CATAAAC	CATAGAC
	6861	CGCTCTG	CTGAAGC	CAGTTAC	CTTCGGA	AAAAGAG	TTGGTAG	CTCTTGA	TCCGGTA	AACAAC	CACCGCT
		GCGAGAC	GACTTCG	GTCAATG	GAAGCCT	TTTTCTC	AACCATC	GAGAACT	AGGCCGT	TTGTTTG	GTGGCGA
	6931	GTGAGCG	GTGGTTT	TTTTGTT	TGCAAGC	AGCAGAT	TACGCGC	AGAAAAA	AAGGATC	TCAAGAA	GATCCTT
		CCATCGC	CACCAAA	AAAACAA	ACGTTCT	TGCTCTA	ATGCGCG	TCTTTT	TTCTTGA	AGTTTAA	TATATAT
	7001	TGATCTT	TTCTACG	GGGTCTG	ACGCTCA	GTGGAAC	GAAAACT	CACGTTA	AGGGATT	TTGTGTA	TGAGATT
		ACTAGAA	AGATATC	CCCAGAC	TGCGAGT	CACCTTG	CTTTTGA	GTGCAAT	TCCCTAA	AACCACT	ACTCTAA
	7071	ATCAAAA	AGATCTT	TCACCTA	GATCCTT	TTAAATG	AAAAATG	AAGTTTT	AAATCAA	TCTAAAG	TATATAT
		TAGTTTT	TCCTAGA	AGTGGAT	CTAGGAA	AATTTAA	TTTTTAC	TTCAAAA	TTTAGTT	AGATTTC	ATATATA

```

7141      GAGTAAA CTTGGTC TGACAGT TACCAAT GCTTAAT CAGTGAG GCACCTA TCTCAGC GATCTGT CTATTTC
          CTCATTT GAACCAG ACTGTCA ATGGTTA CGAATTA GTCACCT CGTGGAT AGAGTCG CTAGACA GATAAAG
          ~~~~~~
5 7211 GTTCATC CATAGTT GCCTGAC TCCCGCT CGTGTAG ATAACCTA CGATACG GGAGGGC TTACCAT CTGGCCC
 CAAGTAG GTATCAA CGGACTG AGGGGCA GCACATC TATTGAT GCTATGC CCFCCCG AATGGTA GACCGGG
 Amp resistance gene
7281 CAGTGCT GCAATGA TACCGCG AGACCCA CGCTCAC CGGCTCC AGATTTA TCAGCAA TAAACCA GCCAGCC
 GTCACGA CGTTACT ATGGCGC TCTGGGT GCGAGTG GCCGAGG TCTAAAT AGTCGTT ATTTGGT CGGTCCG
 Amp resistance gene
10 7351 GGAAGGG CCGAGCG CAGAAGT GGTCTCG CAACCTT ATCCGCC TCCATCC AGTCTAT TAATGTG TGCCGGG
 CCTTCCC GGCTCGC GTCTTCA CCAGGAC GTTGAAA TAGGCGG AGGTAGG TCAGATA ATTAACA ACGGCCC
 Amp resistance gene
15 7421 AAGCTAG AGTAAGT AGTTCGC CAGTTAA TAGTTTG CGCAACG TTGTTGC CATTGCT ACAGGCA TCGTGTG
 TTCGATC TCATTCA TCAAGCG GTCBAAT ATCAAAC GCGTTGC AACCAAC GTAACGA TGTCCGT AGCACCA
 Amp resistance gene
7491 GTCACGC TCGTCGT TTGGTAT GGCTTCA TTCAGCT CCGGTTT CCAACGA TCAAGGC GAGTTAC ATGATCC
 CAGTGCG AGCAGCA AACCATA CCGAAGT AAGTCGA GGCCAAG GGTGCTT AGTTCCG CTCAATG TACTAGG
 Amp resistance gene
20 7561 CCCATGT TGTGCAA AAAAGCG GTTAGCT CCTTCGG TCCTCCG ATCGTTG TCAGAAG TAAGTTG GCCGCAG
 GGGTACA ACACGTT TTTTCGC CAATCGA GGAAGCC AGGAGGC TAGCAAC AGTCTTC ATTC AAC CGGCGTC
 Amp resistance gene
25 7631 TGTATAC ACTCATG GTTATGG CAGCACT GCATAAT TCTCTTA CTGTCAT GCCATCC GTAAGAT GCTTTTC
 ACAATAG TGAGTAC CAATACC GTCGTGA CGTATTA AGAGAAT GACAGTA CCGTAGG CATTCTA CGAAAAG
 Amp resistance gene
7701 TGTGACT GGTGAGT ACTCAAC CAAGTCA TTCTGAG AATAGTG TATGCGG CGACCGA GTTGCTC TTGCCCC
 ACACTGA CCACTCA TGAGTTG GTTCAGT AAGACTC TTATCAC ATACGCC GCTGGCT CAACGAG AACGGGC
 Amp resistance gene
30 7771 GCGTCAA TACGGGA TAATACC GCGCCAC ATAGCAG AACCTTA AAAGTGC TCATCAT TGGAAAA CGTTCCT
 CGCAGTT ATGCCCT ATTATGG CGCGGTG TATCGTC TTGAAAT TTTCACG AGTAGTA ACCTTTT GCAAGAA
 Amp resistance gene
7841 CGGGGCG AAAACTC TCAAGGA TCTTACC GCTGTTG AGATCCA GTTCGAT GTAACCC ACTCGTG CACCCAA
 GCCCGC TTTTGAG AGTTCCT AGAATGG CGACAAC TCTAGGT CAAGCTA CATTGGG TGAGCAC GTGGGTT
 Amp resistance gene
35 7911 CTGATCT TCAGCAT CTTTAC TTTACCC AGCGTTT CTGGGTG AGCAAAA ACAGGAA GGCAAAA TGCCGCA
 GACTAGA AGTCGTA GAAAATG AAAGTGG TCGCAAA GACCCAC TCGTTTT TGTCTTT CCGTTTT ACGGCGT
 Amp resistance gene
7981 AAAAAGG GAATAAG GGCAGCA CGGAAAT GTTGAAT ACTCATA CTCTTCC TTTTCA ATATTAT TGAAGCA
 TTTTTC CTTATTC CCGCTGT GCCTTTA CAACCTA TGAGTAT GAGAAGG AAAAAGT TATAATA ACTTTCGT
          ~~~~~~
40      Amp resistance gene
8051      TTTATCA GGGTTAT TGTCTCA TGAGCGG ATACATA TTTGAAT GTATTTA GAAAAAT AAACAAA TAGGGGT
          AAATAGT CCCAATA ACAGAGT ACTCGCC TATGTAT AAACCTA CATAAAT CTTTTTA TTGTTT ATCCCCA
45 8121      TCCGCGC ACATTTT CCCGAAA AGTGCCA CCTGACG TCTAAGA AACCATT ATTATCA TGACATT AACCTAT
          AGGCGCG TGTAAGG GGGCTTT TCACGGT GGACTGC AGATTCT TTGGTAA TAATAGT ACTGTAA TTGGATA
8191      AAAAATA GCGGTAT CACGAG
          TTTTTAT CCGCATA GTGCTC

```

**FIGURE 2A**

		1				50
	mCEA (6D)	ATGGAGTCTC	CCTCGGCCCC	TCCCCACAGA	TGGTGCATCC	CCTGGCAGAG
5	mCEA (6D, 1st&2nd)	ATGGAGTCTC	CCTCGGCCCC	TCCCCACAGA	TGGTGCATCC	CCTGGCAGAG
		51				100
	mCEA (6D)	GCTCCTGCTC	ACAGCCTCAC	TTCTAACCTT	CTGGAACCCG	CCCACCACTG
	mCEA (6D, 1st&2nd)	GCTCCTGCTC	ACAGCCTCAC	TTCTAACCTT	CTGGAACCCG	CCCACCACTG
10		101				150
	mCEA (6D)	CCAAGCTCAC	TATTGAATCC	ACGCCGTTCA	ATGTCGCAGA	GGGGAAGGAG
	mCEA (6D, 1st&2nd)	CCAAGCTCAC	TATTGAATCC	ACGCCGTTCA	ATGTCGCAGA	GGGGAAGGAG
15		151				200
	mCEA (6D)	GTGCTTCTAC	TTGTCCACAA	TCTGCCCCAG	CATCTTTTTC	GCTACAGCTG
	mCEA (6D, 1st&2nd)	GTGCTTCTAC	TTGTCCACAA	TCTGCCCCAG	CATCTTTTTC	GCTACAGCTG
20		201				250
	mCEA (6D)	GTACAAAGGT	GAAAGAGTGG	ATGGCAACCG	TCAAATTATA	GGATATGTAA
	mCEA (6D, 1st&2nd)	GTACAAAGGT	GAAAGAGTGG	ATGGCAACCG	TCAAATTATA	GGATATGTAA
25		251				300
	mCEA (6D)	TAGGAACTCA	ACAAGCTACC	CCAGGGCCCG	CATACAGTGG	TCGAGAGATA
	mCEA (6D, 1st&2nd)	TAGGAACTCA	ACAAGCTACC	CCAGGGCCCG	CATACAGTGG	TCGAGAGATA
30		301				350
	mCEA (6D)	ATATACCCCA	ATGCATCCCT	GCTGATCCAG	AACATCATCC	AGAATGACAC
	mCEA (6D, 1st&2nd)	ATATACCCCA	ATGCATCCCT	GCTGATCCAG	AACATCATCC	AGAATGACAC
35		351				400
	mCEA (6D)	AGGATTCTAC	ACCCTACACG	TCATAAAGTC	AGATCTTGTG	AATGAAGAAG
	mCEA (6D, 1st&2nd)	AGGATTCTAC	ACCCTACACG	TCATAAAGTC	AGATCTTGTG	AATGAAGAAG
40		401				450
	mCEA (6D)	CAACTGGCCA	GTTCCGGGTA	TACCCGGAGC	TGCCCAAGCC	CTCCATCTCC
	mCEA (6D, 1st&2nd)	CAACTGGCCA	GTTCCGGGTA	TACCCGGAAC	TCCCTAAGCC	TTCTATTAGC
45		451				500
	mCEA (6D)	AGCAACAAC	CCAAACCCGT	GGAGGACAAG	GATGCTGTGG	CCTTCACCTG
	mCEA (6D, 1st&2nd)	<u>TC</u> CAATAATA	<u>G</u> TAAGCCTGT	<u>C</u> GAAGACAAA	<u>G</u> ATGCCGT <u>C</u> G	<u>C</u> TTT <u>T</u> ACATG
50		501				550
	mCEA (6D)	TGAACCTGAG	ACTCAGGACG	CAACCTACCT	GTGGTGGGTA	AACAATCAGA
	mCEA (6D, 1st&2nd)	<u>C</u> GAGCC <u>C</u> GAA	ACTCAAGACG	CAACATATCT	<u>C</u> TGGTGGG <u>T</u> G	AACAAC <u>C</u> AGT
55		551				600
	mCEA (6D)	GCCTCCCGGT	CAGTCCCAGG	CTGCAGCTGT	CCAATGGCAA	CAGGACCCTC
	mCEA (6D, 1st&2nd)	<u>C</u> CTG <u>C</u> CTGT	<u>G</u> TCCCCTAG <u>A</u>	<u>C</u> TCCA <u>A</u> CTCA	<u>G</u> CAACGG <u>A</u> AA	<u>T</u> AGAACTCTG
60		601				650
	mCEA (6D)	ACTCTATTCA	ATGTCACAAG	AAATGACACA	GCAAGCTACA	AATGTGAAAC
	mCEA (6D, 1st&2nd)	ACCCTGTTTA	ACGTGACCAG	GAACGACACA	GCAAGCTACA	AATG <u>C</u> GAAAC

**FIGURE 2B**

		651				700
	mCEA (6D)	CCAGA <u>AC</u> CCA	GTGAGTGCCA	GGCGCAGTGA	TTCAGTCATC	CTGAATGTCC
5	mCEA (6D, 1st&2nd)	CCAAA <u>AT</u> CCA	GTCAGCGCCA	GGAGG <u>T</u> CTGA	TTCAGTGATT	CTCAAC <u>G</u> TGC
		701				750
	mCEA (6D)	TCTATGGCCC	GGATGCCCC	ACCATTTCCC	CTCTAAACAC	ATCTTACAGA
	mCEA (6D, 1st&2nd)	T <u>T</u> TACGGACC	CGATGCTCCT	ACAAT <u>CAG</u> CC	CTCTAAACAC	AAGCTAT <u>AG</u> A
10		751				800
	mCEA (6D)	TCAGGGGAAA	ATCTGAACCT	CTCCTGCCAC	GCAGCCTCTA	ACCCACCTGC
	mCEA (6D, 1st&2nd)	TCAGGGGAAA	ATCTGAATCT	<u>GAG</u> CTGTCAT	GCCGCTAGCA	ATCCTCCCGC
15		801				850
	mCEA (6D)	ACAGTACTCT	TGGTTTGTCA	ATGGGACTTT	CCAGCAATCC	ACCCAAGAGC
	mCEA (6D, 1st&2nd)	<u>CCA</u> ATACAGC	TGGTTTGTCA	ATGGC <u>ACT</u> TT	CCAACAGTCC	ACCCAGGA <u>C</u>
20		851				900
	mCEA (6D)	TCTTTATCCC	CAACATCACT	GTGAATAATA	GTGGATCCTA	TACGTGCCAA
	mCEA (6D, 1st&2nd)	TGTT <u>CAT</u> TCC	CAATATTACC	GTGAACAATA	GTGGATCCTA	<u>CAC</u> GTGCCAA
25		901				950
	mCEA (6D)	GCCCATAACT	CAGACACTGG	CCTCAATAGG	ACCACAGTCA	CGACGATCAC
	mCEA (6D, 1st&2nd)	GCT <u>CACA</u> ATA	<u>GCG</u> ACACCGG	<u>ACT</u> CAACCGC	ACAACCGTGA	CGACGATTAC
30		951				1000
	mCEA (6D)	AGTCTATGAG	CCACCCAAAC	CCTTCATCAC	CAGCAACAAC	TCCAACCCCG
	mCEA (6D, 1st&2nd)	<u>CGT</u> GTATGAG	CCACCAAAAC	CATT <u>CATA</u> AC	<u>TAG</u> TAACAAT	TCTAACC <u>CAG</u>
35		1001				1050
	mCEA (6D)	TGGAGGATGA	GGATGCTGTA	GCCTTAACCT	GTGAACCTGA	GATTTCAGAAC
	mCEA (6D, 1st&2nd)	T <u>TG</u> AGGATGA	GGACGCAGTT	GCATTAACTT	GTGAGCCAGA	GATT <u>CAAA</u> T
40		1051				1100
	mCEA (6D)	ACAACCTACC	TGTGGTGGGT	AAATAATCAG	AGCCTCCCGG	TCAGTCCCAG
	mCEA (6D, 1st&2nd)	ACC <u>ACT</u> TATT	TATGGTGGGT	<u>CA</u> ATAACCAA	AGTTTGCCGG	TTAGCC <u>ACG</u>
45		1101				1150
	mCEA (6D)	GCTGCAGCTG	TCCAATGACA	ACAGGACCCCT	CACTCTACTC	AGTGTACAA
	mCEA (6D, 1st&2nd)	<u>CTT</u> GCAGTTG	TCTAATGATA	ACCGCACATT	<u>GAC</u> ACTCCTG	<u>TCC</u> GTTACTC
50		1151				1200
	mCEA (6D)	GGAATGATGT	AGGACCCCTAT	GAGTGTGGAA	TCCAGAACGA	ATTAAGTGTT
	mCEA (6D, 1st&2nd)	GCAATGATGT	AGGACCTTAT	GAGTGTGGCA	TTCAGAA <u>TGA</u>	ATTATCCGTT
55		1201				1250
	mCEA (6D)	GACCACAGCG	ACCCAGTCAT	CCTGAATGTC	CTCTATGGCC	CAGACGACCC
	mCEA (6D, 1st&2nd)	GATC <u>ACT</u> CCG	ACCCTGTTAT	CCTTAATGTT	<u>TTG</u> TATGGCC	CAGACGACCC
60		1251				1300
	mCEA (6D)	CACCA <u>TTT</u> CC	CCCTCATACA	CCTATTACCG	TCCAGGGGTG	AACCTCAGCC
	mCEA (6D, 1st&2nd)	<u>AACT</u> ATATCT	CCATCATACA	CCTACTACCG	TCCGGGCGTG	AACTT <u>GAG</u> CC



**FIGURE 2C**

		1301				1350
	mCEA (6D)	TCTCCTGCCA	TGCAGCCTCT	AACCCACCTG	CACAGTATTC	TTGGCTGATT
5	mCEA (6D, 1st&2nd)	<u>TTTCT</u> TGCCA	TGCAGC <u>ATCC</u>	AACCC <u>CCCTG</u>	CACAGT <u>ATTC</u>	<u>CTGGCTGATT</u>
		1351				1400
	mCEA (6D)	GATGGGAACA	TCCAGCAACA	CACACAAGAG	CTCTTTATCT	CCAACATCAC
	mCEA (6D, 1st&2nd)	GATGG <u>AAACA</u>	T <u>TCAGCAGCA</u>	<u>TACTCAAGAG</u>	<u>TTATTTATAA</u>	<u>GCAACATAAC</u>
10		1401				1450
	mCEA (6D)	TGAGAAGAAC	AGCGGACTCT	ATACCTGCCA	GGCCAATAAC	TCAGCCAGTG
	mCEA (6D, 1st&2nd)	TGAGAAGAAC	AGCGGACTCT	ATACT <u>TGCCA</u>	GGCCAATAAC	TCAGCCAGTG
15		1451				1500
	mCEA (6D)	GCCACAGCAG	GACTACAGTC	AAGACAATCA	CAGTCTCTGC	GGAGCTGCCC
	mCEA (6D, 1st&2nd)	<u>GTCACAGCAG</u>	GACTACAG <u>TT</u>	<u>AAAACAATAA</u>	<u>CTGTTC</u> <u>CCGC</u>	GGAGCTGCCC
20		1501				1550
	mCEA (6D)	AAGCCCTCCA	TCTCCAGCAA	CAACTCCAAA	CCCGTGGAGG	ACAAGGATGC
	mCEA (6D, 1st&2nd)	AAGCCCTCCA	TCTCCAGCAA	CAACTCCAAA	CCCGTGGAGG	ACAAGGATGC
25		1551				1600
	mCEA (6D)	TGTGGCCTTC	ACCTGTGAAC	CTGAGGCTCA	GAACACAACC	TACCTGTGGT
	mCEA (6D, 1st&2nd)	TGTGGCCTTC	ACCTGTGAAC	CTGAGGCTCA	GAACACAACC	TACCTGTGGT
30		1601				1650
	mCEA (6D)	GGGTAAATGG	TCAGAGCCTC	CCAGTCAGTC	CCAGGCTGCA	GCTGTCCAAT
	mCEA (6D, 1st&2nd)	GGGTAAATGG	TCAGAGCCTC	CCAGTCAGTC	CCAGGCTGCA	GCTGTCCAAT
35		1651				1700
	mCEA (6D)	GGCAACAGGA	CCCTCACTCT	ATTCAATGTC	ACAAGAAATG	ACGCAAGAGC
	mCEA (6D, 1st&2nd)	GGCAACAGGA	CCCTCACTCT	ATTCAATGTC	ACAAGAAATG	ACGCAAGAGC
40		1701				1750
	mCEA (6D)	CTATGTATGT	GGAATCCAGA	ACTCAGTGAG	TGCAAACCGC	AGTGACCCAG
	mCEA (6D, 1st&2nd)	CTATGTATGT	GGAATCCAGA	ACTCAGTGAG	TGCAAACCGC	AGTGACCCAG
45		1751				1800
	mCEA (6D)	TCACCCCTGGA	TGTCCTCTAT	GGGCCGGACA	CCCCATCAT	TTCCCCCCCCA
	mCEA (6D, 1st&2nd)	TCACCCCTGGA	TGTCCTCTAT	GGGCCGGACA	CCCCATCAT	TTCCCCCCCCA
50		1801				1850
	mCEA (6D)	GACTCGTCTT	ACCTTTCGGG	AGCGGACCTC	AACCTCTCCT	GCCACTCGGC
	mCEA (6D, 1st&2nd)	GACTCGTCTT	ACCTTTCGGG	AGCGGACCTC	AACCTCTCCT	GCCACTCGGC
		1851				1900
	mCEA (6D)	CTCTAACCCA	TCCCCGCAGT	ATTCTTGCG	TATCAATGGG	ATACCGCAGC
	mCEA (6D, 1st&2nd)	CTCTAACCCA	TCCCCGCAGT	ATTCTTGCG	TATCAATGGG	ATACCGCAGC
		1901				1950
	mCEA (6D)	AACACACACA	AGTTCTCTTT	ATCGCCAAAA	TCACGCCAAA	TAATAACGGG
	mCEA (6D, 1st&2nd)	AACACACACA	AGTTCTCTTT	ATCGCCAAAA	TCACGCCAAA	TAATAACGGG

**FIGURE 2D**

		1951				2000
	mCEA (6D)	ACCTATGCCT	GTTTTGTCTC	TAACTTGGCT	ACTGGCCGCA	ATAATTCCAT
5	mCEA (6D, 1st&2nd)	ACCTATGCCT	GTTTTGTCTC	TAACTTGGCT	ACTGGCCGCA	ATAATTCCAT
		2001				2050
	mCEA (6D)	AGTCAAGAGC	ATCACAGTCT	CTGCATCTGG	AACTTCTCCT	GGTCTCTCAG
10	mCEA (6D, 1st&2nd)	AGTCAAGAGC	ATCACAGTCT	CTGCATCTGG	AACTTCTCCT	GGTCTCTCAG
		2051				2100
	mCEA (6D)	CTGGGGCCAC	TGTCGGCATC	ATGATTGGAG	TGCTGGTTGG	GGTTGCTCTG
	mCEA (6D, 1st&2nd)	CTGGGGCCAC	TGTCGGCATC	ATGATTGGAG	TGCTGGTTGG	GGTTGCTCTG
15		2101				
	mCEA (6D)	ATATAG				
	mCEA (6D, 1st&2nd)	ATATAG				

**FIGURE 3****A. Amino Acid Sequence Comparison of "Wild-Type KSA" (1) and Modified KSA (2)**

5 1 MAPPQVLAFLGLLLAAATATFAAAQEEVCENYKLAVNCFVNNNRQCQCTSVGAQNTVIC  
 2 MAPPQVLAFLGLLLAAATATFAAAQEEVCENYKLAVNCFVNNNRQCQCTSVGAQNTVIC

1 SKLAAKCLVMKAEMNGSKLGRRAKPEGALQNNNDGLYDPDCDESGLFKAKQCNGTSTCWC  
 2 SKLAAKCLVMKAEMNGSKLGRRAKPEGALQNNNDGLYDPDCDESGLFKAKQCNGTSTCWC

10 1 VNTAGVRRTDKDTEITCSEVRITYWIIIELEKHKAREKPYDSKSLRTALQKEITTRYQLD  
 2 VNTAGVRRTDKDTEITCSEVRITYWIIIELEKHKAREKPYDSKSLRTALQKEITTRYQLD

1 PKFITSILYENNVITIDLQNSSQKTQNDVDIADVAYYFEKDVKGESLFHSHKMDLTVN  
 15 2 PKFITSILYENNVITIDLQNSSQKTQNDVDIADVAYYFEKDVKGESLFHSHKMDLTVN

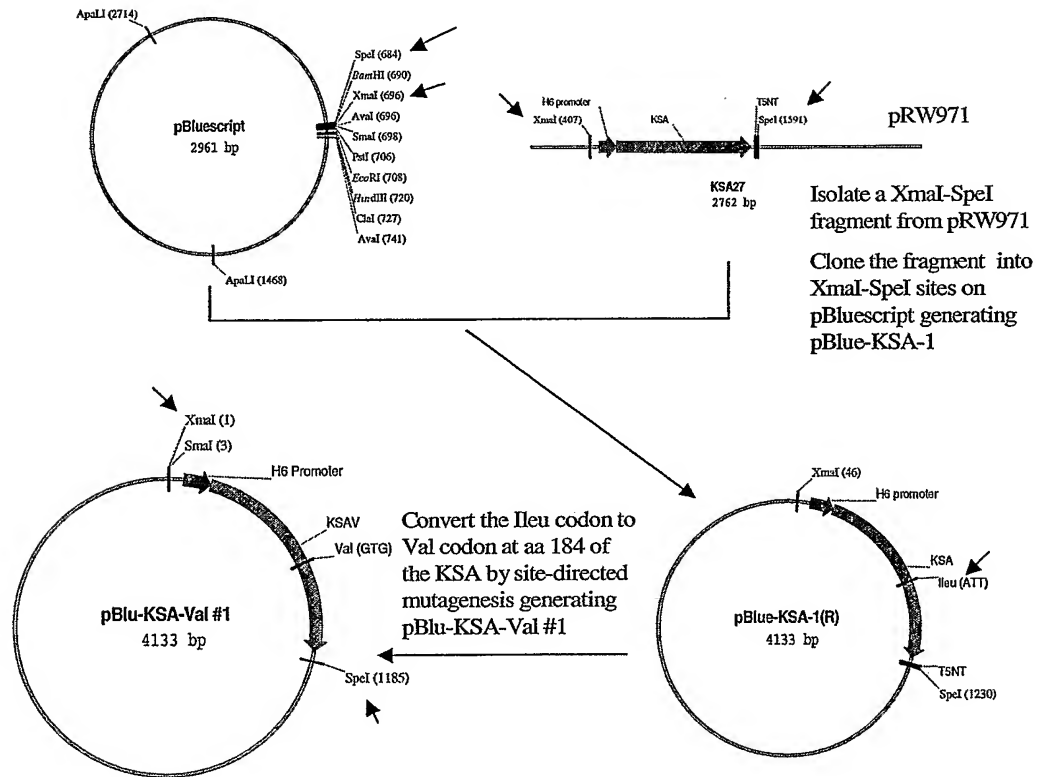
1 GEQLDLDPGQTLIIYYVDEKAPEFSMQGLKAGVIAVIVVVVIAVVAGIVVLVISRKKRMA  
 2 GEQLDLDPGQTLIIYYVDEKAPEFSMQGLKAGVIAVIVVVVIAVVAGIVVLVISRKKRMA

20 1 KYEKAETIKEMGEMHRELNA  
 2 KYEKAETIKEMGEMHRELNA

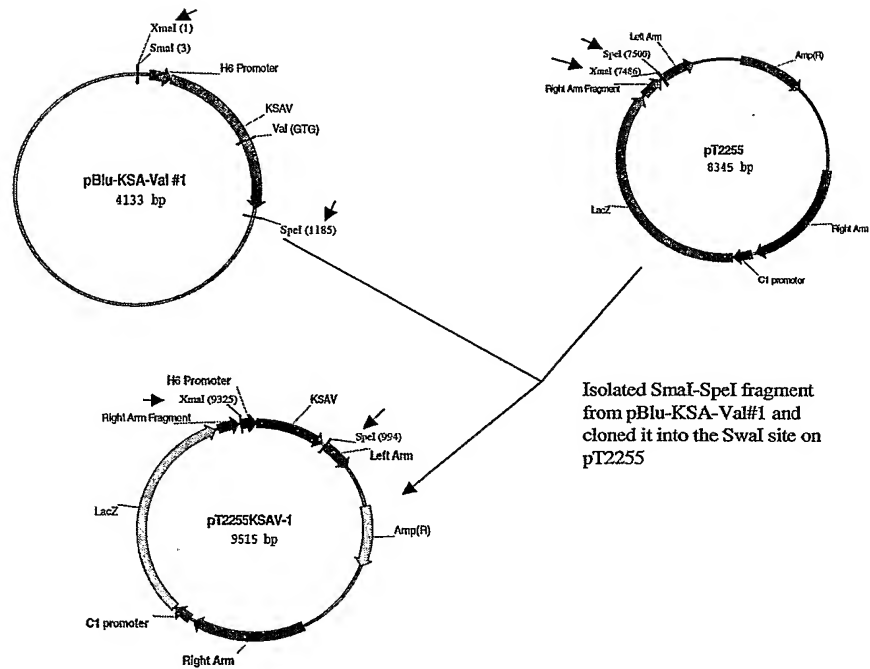
**B. DNA Sequence of Modified KSA**

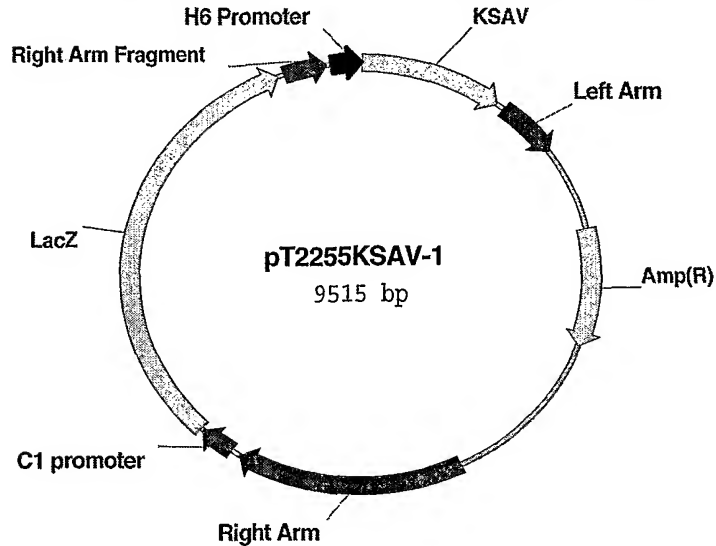
atggcgcccccgaggctcctcgcggttcgggcttctgcttgccgcgggcgacggcgacttttgccgcagctcaggaa  
 25 gaatgtgtctgtgaaaactacaagctggccgtaaactgctttgtgaataataatcgatcaatgccagtgacttca  
 gttggtgcacaaaatactgtcatttgctcaaagctggctgccaaatgcttggtgatgaaggcagaaatgaatggc  
 tcaaaacttgggagaagagcaaaacctgaagggggccctccagaacaatgatgggctttatgatcctgactgcgat  
 gagagcgggctctttaaggccaagcagtgcaacggcacctccacgtgctgggtgtgtgaactgctgggggtcaga  
 agaacagacaaggacactgaaataacctgctctgagcagtgagaacctactggatcatcattgaactaaaacac  
 30 aaagcaagagaaaaaccttatgatagtaaaagtttgcgactgcacttcagaaggagatcacaacgcgttatcaa  
 ctggatccaaaatttatcacgagtggtgtgtatgagaataatgttatcactattgatctggttcaaaattcttct  
 caaaaaactcagaatgatgtggacatagctgatgtggcttattattttgaaaagatgttaaagggtgaatccttg  
 ttctattcctaagaaaatggacctgacagtaaatggggaacaactggatctggatcctggtcaaaactttaatttat  
 tatgttgatgaaaaagcacctgaattctcaatgcagggtctaaaagctgggtgttattgctgttattgtggtgtg  
 35 gtgatagcagttgttgctggaattgtgtgctgggtatttccagaaagaagagaatggcaaagtatgagaaggct  
 gagataaaggagatgggtgagatgcatagggaactcaatgcataa

**FIGURE 4A**  
**Construction of Modified KSA Plasmid**



**FIGURE 4B**  
**Construction of Modified KSA Plasmid**



**FIGURE 5****A. Plasmid Map of Modified KSA Expression Vector****B. DNA Sequence of Modified KSA Expression Vector**

<b>Promoter H6 for KSAV</b>	<b>9930-9515</b>
<b>KSAV</b>	<b>1-945</b>
<b>Left arm</b>	<b>1002-1422</b>
<b>Right arm</b>	<b>4070-5590</b>
<b>Right arm fragment</b>	<b>9012-9299</b>

1 MetAlaProPro GlnValLeu AlaPheGly LeuLeuLeuAla AlaAlaThr-  
 10 ATGGCGCCCC CGCAGGTCCT CGCGTTCGGG CTTCTGCTTG CCGCGGCGAC  
 TACCGCGGGG GCGTCCAGGA GCGCAAGCCC GAAGACGAAC GGCGCCGCTG  
 .AlaThrPhe AlaAlaAlaGln GluGluCys ValCysGlu AsnTyrLysLeu-  
 51 GGCGACTTTT GCCGAGCTC AGGAAGAATG TGTCTGTGAA AACTACAAGC  
 CCGCTGAAAA CGGCGTCGAG TCCTTCTTAC ACAGACACTT TTGATGTTCTG  
 ..AlaValAsn CysPheVal AsnAsnAsnArg GlnCysGln CysThrSer  
 15 101 TGGCCGTAAA CTGCTTTGTG AATAATAATC GTCAATGCCA GTGTACTTCA  
 ACCGGCATTG GACGAAACAC TTATTATTAG CAGTTACGGT CACATGAAGT  
 ValGlyAlaGln AsnThrVal IleCysSer LysLeuAlaAla LysCysLeu-  
 151 GTTGGTGAC AAAATACTGT CATTGCTCA AAGCTGGCTG CCAAATGTTT  
 CAACCACGTG TTTTATGACA GTAAACGAGT TTCGACCGAC GGTTTACAAA  
 20 .ValMetLys AlaGluMetAsn GlySerLys LeuGlyArg ArgAlaLysPro-  
 201 GGTGATGAAG GCAGAAATGA ATGGCTCAA ACTTGGGAGA AGAGCAAAAC  
 CCACTACTTC CGTCTTTACT TACCGAGTTT TGAACCCTCT TCTCGTTTGT  
 ..GluGlyAla LeuGlnAsn AsnAspGlyLeu TyrAspPro AspCysAsp  
 25 251 CTGAAGGGGC CCTCCAGAAC AATGATGGG TTTATGATCC TGACTGCGAT  
 GACTTCCCCG GGAGGTCTTG TTACTACCCG AAATACTAGG ACTGACGCTA  
 GluSerGlyLeu PheLysAla LysGlnCys AsnGlyThrSer ThrCysTrp-  
 301 GAGAGCGGGC TCTTTAAGGC CAAGCAGTGC AACGGCACCT CCACGTGCTG  
 CTCTCGCCCG AGAAATTCCG GTTCGTCACG TTGCCGTGGA GGTGCACGAC  
 .CysValAsn ThrAlaGlyVal ArgArgThr AspLysAsp ThrGluIleThr-  
 30 351 GTGTGTGAAC ACTGCTGGGG TCAGAAGAAC AGACAAGGAC ACTGAAATAA  
 CACACACTTG TGACGACCCC AGTCTTCTTG TCTGTTCTCTG TGACTTTATT  
 ..CysSerGlu ArgValArg ThrTyrTrpIle IleIleGlu LeuLysHis

401 CCTGCTCTGA GCGAGTGAGA ACCTACTGGA TCATCATTGA ACTAAAACAC  
 GGACGAGACT CGCTCACTCT TGGATGACCT AGTAGTAACT TGATTTTGTG  
 LysAlaArgGlu LysProTyr AspSerLys SerLeuArgThr AlaLeuGln.  
 451 AAAGCAAGAG AAAAACCTTA TGATAGTAAA AGTTTGCGGA CTGCACCTCA  
 5 TTTTCGTTCTC TTTTGGGAAT ACTATCATTT TCAAACGCCT GACGTGAAGT  
 .LysGluIle ThrThrArgTyr GlnLeuAsp ProLysPhe IleThrSerVal.  
 501 GAAGGAGATC ACAACGCGTT ATCAACTGGA TCCAAAATTT ATCACGAGTG  
 CTTCCTCTAG TGTGCGCAA TAGTTGACCT AGGTTTTAAA TAGTGCTCAC  
 ..LeuTyrGlu AsnAsnVal IleThrIleAsp LeuValGln AsnSerSer  
 10 551 TGTGTATGA GAATAATGTT ATCACTATTG ATCTGGTTCA AAATTCCTCT  
 ACAACATACT CTTATTACAA TAGTGATAAC TAGACCAAGT TTTAAGAAGA  
 GlnLysThrGln AsnAspVal AspIleAla AspValAlaTyr TyrPheGlu.  
 601 CAAAAAAGT AGAATGATGT GGACATAGCT GATGTGGCTT ATTATTTTGA  
 GTTTTTTCTG TCTTACTACA CCTGTATCGA CTACACCGAA TAATAAAACT  
 15 .LysAspVal LysGlyGluSer LeuPheHis SerLysLys MetAspLeuThr.  
 651 AAAAGATGTT AAAGGTGAAT CCTTGTTC TTTCTAAGAAA ATGGACCTGA  
 TTTTCTACAA TTTCCACTTA GGAACAAAGT AAGATTCTTT TACCTGGACT  
 ..ValAsnGly GluGlnLeu AspLeuAspPro GlyGlnThr LeuIleTyr  
 701 CAGTAAATGG GGAACAATG GATCTGGATC CTGGTCAAAC TTTAATTTAT  
 20 GTCAATTACG CCTTGTGAC CTAGACCTAG GACCAGTTT AAATTAAATA  
 TyrValAspGlu LysAlaPro GluPheSer MetGlnGlyLeu LysAlaGly.  
 751 TATGTTGATG AAAAAGCACC TGAATTCTCA ATGCAGGGTC TAAAAGCTGG  
 ATCAACTAC TTTTTCGTGG ACTTAAGAGT TACGTCCCAG ATTTTCGACC  
 .ValIleAla ValIleValVal ValValIle AlaValVal AlaGlyIleVal.  
 25 801 TGTATTGCT GTTATTGTGG TTGTGGTGAT AGCAGTTGTT GCTGGAATTG  
 ACAATAACGA CAATAACACC AACACCACTA TCGTCAACAA CGACCTTAAC  
 ..ValLeuVal IleSerArg LysLysArgMet AlaLysTyr GluLysAla  
 851 TTGTGCTGGT TATTTCCAGA AAGAAGAGAA TGGCAAAGTA TGAGAAGGCT  
 AACACGACCA ATAAAGGTCT TTCTTCTCTT ACCGTTTCAT ACTCTCCGA  
 30 GluIleLysGlu MetGlyGlu MetHisArg GluLeuAsnAla ***  
 901 GAGATAAAGG AGATGGGTGA GATGCATAGG GAACTCAATG CATAAGAAGC  
 CTCTATTTCC TCTACCCACT CTACGTATCC CTGAGTTAC GTATTCTTCG  
 951 TTATCGATAC CGTCGACCTC GAGGAATTCT TTTTATTGAT TAACTAGTTA  
 AATAGCTATG GCAGCTGGAG CTCTTAAGA AAAATAACTA ATTGATCAAT  
 35 1001 ATCACGCGCG CTTATAAAGA TCTAAATGC ATAATTTCTA AATAATGAAA  
 TAGTGCCGGC GAATATTTCT AGATTTTACG TATTAAAGAT TTATTACTTT  
 1051 AAAAAGTACA TCATGAGCAA CGCGTTAGTA TATTTTACAA TGGAGATTAA  
 TTTTTCATGT AGTACTCGTT GCGCAATCAT ATAAAAATGTT ACCTCTAATT  
 1101 CGCTCTATAC CGTTCTATGT TTATTGATTC AGATGATGTT TTAGAAAAGA  
 40 GCGAGATATG GCAAGATACA AATAACTAAG TCTACTACAA AATCTTTTCT  
 1151 AAGTTATTGA ATATGAAAAC TTTAATGAAG ATGAAGATGA CGACGATGAT  
 TTCAATAACT TATACTTTTG AAATTACTTC TACTTCTACT GCTGCTACTA  
 1201 TATTGTTGTA AATCTGTTTT AGATGAAGAA GATGACGCGC TAAAGTATAC  
 ATAACAACAT TTAGACAAA TCTACTTCTT CTACTGCGCG ATTTTCATATG  
 45 1251 TATGGTTACA AAGTATAAGT CTATACTACT AATGGCGACT TGTGCAAGAA  
 ATACCAATGT TTCATATTCA GATATGATGA TTACCGCTGA ACACGTTCTT  
 1301 GGTATAGTAT AGTGAAAATG TTGTTAGATT ATGATTATGA AAAACCAAAT  
 CCATATCATA TCACTTTTAC AACAATCTAA TACTAATACT TTTTGGTTTA  
 1351 AAATCAGATC CATATCTAAA GGTATCTCCT TTGCACATAA TTTTCATCTAT  
 50 TTTAGTCTAG GTATAGATTT CCATAGAGGA AACGTGTATT AAAGTAGATA  
 1401 TCCTAGTTTA GAATACCTGC AGCCAAGCTT GGCACTGGCC GTCGTTTTAC  
 AGGATCAAAT CTTATGGACG TCGGTTTCGAA CCGTGACCGG CAGCAAAATG  
 1451 AACGTCGTGA CTGGGAAAAC CCTGGCGTTA CCCAACTTAA TCGCCTTGCA  
 TTGCAGCACT GACCCTTTTG GGACCGCAAT GGGTTGAATT AGCGGAACGT  
 55 1501 GCACATCCCC CTTTCGCCAG CTGGCGTAAT AGCGAAGAGG CCCGACCGA  
 CGTGTAGGGG GAAAGCGGTC GACCGCATTA TCGCTTCTCC GGGCGTGGCT  
 1551 TCGCCCTTCC CAACAGTTGC GCAGCCTGAA TGGCGAATGG CGCCTGATGC  
 AGCGGGAAGG GTTGTCAACG CGTCGGACTT ACCGCTTACC GCGGACTACG  
 1601 GGTATTTTCT CCTTACGCAT CTGTGCGGTA TTTCACACCG CATATGGTGC

		CCATAAAAGA	GGAATGCGTA	GACACGCCAT	AAAGTGTGGC	GTATACCACG
1651		ACTCTCAGTA	CAATCTGCTC	TGATGCCGCA	TAGTTAAGCC	AGCCCCGACA
		TGAGAGTCAT	GTTAGACGAG	ACTACGGCGT	ATCAATTCGG	TCGGGGCTGT
5	1701	CCCGCCAACA	CCCGCTGACG	CGCCCTGACG	GGCTTGTCTG	CTCCCGGCAT
		GGGCGGTTGT	GGGCGACTGC	GCGGGACTGC	CCGAACAGAC	GAGGGCCGTA
	1751	CCGCTTACAG	ACAAGCTGTG	ACCGTCTCCG	GGAGCTGCAT	GTGTCAGAGG
		AGCGAATGTC	TGTTTCGACAC	TGGCAGAGGC	CCTCGACGTA	CACAGTCTCC
	1801	TTTTCACCGT	CATCACCGAA	ACGCGCGAGA	CGAAAGGGCC	TCGTGATACG
		AAAAGTGGCA	GTAGTGGCTT	TGCGCGCTCT	GCTTTCCCGG	AGCACTATGC
10	1851	CCTATTTTTA	TAGGTAAATG	TCATGATAAT	AATGGTTTCT	TAGACGTCAG
		GGATAAAAAAT	ATCCAATTAC	AGTACTATTA	TTACCAAAGA	ATCTGCAGTC
	1901	GTGGCACTTT	TCGGGGAAAT	GTGCGCGGAA	CCCCTATTTG	TTTATTTTTT
		CACCGTGAAA	AGCCCCTTTA	CACGCGCCTT	GGGGATAAAC	AAATAAAAAAG
	1951	TAAATACATT	CAAATATGTA	TCCGCTCATG	AGACAATAAC	CCTGATAAAT
15		ATTTATGTAA	GTTTATACAT	AGGCGAGTAC	TCTGTTATTG	GGACTATTTA
	2001	GCTTCAATAA	TATTGAAAAA	GGAAGAGTAT	GAGTATTCAA	CATTTCCGTG
		CGAAGTTATT	ATAACTTTTT	CCTTCTCATA	CTCATAAGTT	GTAAAGGCAC
	2051	TCGCCCTTAT	TCCCTTTTTT	GCGGCATTTT	GCCTTCCTGT	TTTTTGCTCAC
		AGCGGGAATA	AGGGAAAAAA	CGCCGTAAAA	CGGAAGGACA	AAAACGAGTG
20	2101	CCAGAAAACGC	TGGTGAAAGT	AAAAGATGCT	GAAGATCAGT	TGGGTGCACG
		GGTCTTTGCG	ACCACTTTCA	TTTTCTACGA	CTTCTAGTCA	ACCCACGTGC
	2151	AGTGGGTTAC	ATCGAACTGG	ATCTCAACAG	CGGTAAGATC	CTTGAGAGTT
		TCACCCAATG	TAGCTTGACC	TAGAGTTGTC	GCCATTCTAG	GAACCTCTCA
	2201	TTGCCCCCGA	AGAACGTTTT	CCAATGATGA	GCACTTTTAA	AGTTCTGCTA
25		AAGCGGGGCT	TCTTGCAAAA	GGTTACTACT	CGTGAAAATT	TCAAGACGAT
	2251	TGTGGCGCGG	TATTATCCCG	TATTGACGCC	GGGCAAGAGC	AACTCGGTCTG
		ACACCGCGCC	ATAATAGGGC	ATAACTGCGG	CCCCTTCTCG	TTGAGCCAGC
	2301	CCGCATACAC	TATTCTCAGA	ATGACTTGGT	TGAGTACTCA	CCAGTCACAG
		GGCGTATGTG	ATAAGAGTCT	TACTGAACCA	ACTCATGAGT	GGTCAGTGTG
30	2351	AAAAGCATCT	TACGGATGGC	ATGACAGTAA	GAGAATTATG	CAGTGCTGCC
		TTTTCGTAGA	ATGCCTACCG	TACTGTCTAT	CTCTTAATAC	GTACACGACG
	2401	ATAACCATGA	GTGATAACAC	TGCGGCCAAC	TTACTTCTGA	CAACGATCGG
		TATTGGTACT	CACATATTGTG	ACGCCGGTTG	AATGAAGACT	GTTGCTAGCC
	2451	AGGACCGAAG	GAGCTAACCG	CTTTTTTGCA	CAACATGGGG	GATCATGTAA
35		TCCTGGCTTC	CTCGATTGGC	GAAAAACGT	GTTGTACCCC	CTAGTACATT
	2501	CTCGCCTTGA	TCGTTGGGAA	CCGGAGCTGA	ATGAAGCCAT	ACCAAACGAC
		GAGCGGAAC	AGCAACCCTT	GGCCTCGACT	TACTTCGGTA	TGGTTTGCTG
	2551	GAGCGTGACA	CCACGATGCC	TGTAGCAATG	GCAACAACGT	TGCGCAAAC
		CTCGCACTGT	GGTGCTACGG	ACATCGTTAC	CGTTGTTGCA	ACGCGTTTGA
40	2601	ATTAACTGGC	GAACACTTAA	CTCTAGCTTC	CCGGCAACAA	TTAATAGACT
		TAATTGACCG	CTTGATGAAT	GAGATCGAAG	GGCCGTTGTT	AATTATCTGA
	2651	GGATGGAGGC	GGATAAAGTT	GCAGGACCAC	TTCTGCGCTC	GGCCCTTCCG
		CCTACCTCCG	CCTATTTCAA	CGTCCTGGTG	AAGACGCGAG	CCGGGAAGGC
	2701	GCTGGCTGGT	TTATTGCTGA	TAAATCTGGA	GCCGGTGAGC	GTGGGTCTCG
45		CGACCGACCA	AATAACGACT	ATTTAGACCT	CGGCCACTCG	CACCCAGAGC
	2751	CGGTATCATT	GCAGCACTGG	GGCCAGATGG	TAAGCCCTCC	CGTATCGTAG
		GCCATAGTAA	CGTCGTGACC	CCGGTCTACC	ATTCCGGGAG	GCATAGCATC
	2801	TTATCTACAC	GACGGGGAGT	CAGGCAACTA	TGGATGAACG	AAATAGACAG
		AATAGATGTG	CTGCCCCCTCA	GTCCGTTGAT	ACCTACTTGC	TTTATCTGTC
50	2851	ATCGCTGAGA	TAGGTGCCTC	ACTGATTAAG	CATTGGTAAC	TGTCAGACCA
		TAGCGACTCT	ATCCACGGAG	TGACTAATTC	GTAACCATTG	ACAGTCTGGT
	2901	AGTTTACTCA	TATATACTTT	AGATTGATTT	AAAACTTCAT	TTTTAATTTA
		TCAAAATGAGT	ATATATGAAA	TCTAACTAAA	TTTTGAAGTA	AAAATTAAT
	2951	AAAGGATCTA	GGTGAAGATA	CTTTTTTGATA	ATCTCATGAC	CAAAATCCCT
55		TTTCCTAGAT	CCACTTCTAG	GAAAAACTAT	TAGAGTACTG	GTTTTAGGGA
	3001	TAACGTGAGT	TTTCGTTCCA	CTGAGCGTCA	GACCCCGTAG	AAAAGATCAA
		ATTGCACTCA	AAAGCAAGGT	GACTCGCAGT	CTGGGGCATC	TTTTCTAGTT
	3051	AGGATCTTCT	TGAGATCCTT	TTTTTCTGCG	CGTAATCTGC	TGCTTGCAAA
		TCCTAGAAGA	ACTCTAGGAA	AAAAAGACGC	GCATTAGACG	ACGAACGTTT



3101	CAAAAAAACC	ACCGCTACCA	GCGGTGGTTT	GTTTGCCGGA	TCAAGAGCTA
	GTTTTTTTTG	TGGCGATGGT	CGCCACCAAA	CAAACGGCCT	AGTTCCTCGAT
3151	CCAACTCTTT	TTCCGAAGGT	AACTGGCTTC	AGCAGAGCGC	AGATACCAAA
	GGTTGAGAAA	AAGGCTTCCA	TTGACCGAAG	TCGTCTCGCG	TCTATGGTTT
5	3201	TACTGTCCTT	CTAGTGTAGC	CGTAGTTAGG	CCACCACTTC
		ATGACAGGAA	GATCACATCG	GCATCAATCC	GGTGGTGAAG
	3251	TAGCACCGCC	TACATACCTC	GCTCTGCTAA	TCCTGTTACC
		ATCGTGGCGG	ATGTATGGAG	CGAGACGATT	AGGACAATGG
	3301	GCCAGTGGCG	ATAAGTCGTG	TCTTACCGGG	TTGGACTCAA
10		CGGTCACCGC	TATTTCAGCAC	AGAATGGCCC	AACCTGAGTT
	3351	ACCGGATAAG	GCGCAGCGGT	CGGGCTGAAC	GGGGGGTTTC
		TGGCCTATTG	CGCGTCGCCA	GCCCCACTTG	CCCCCAAGC
	3401	CCAGCTTGGA	GCGAACGACC	TACACCGAAC	TGAGATACCT
		GGTCGAACCT	CGCTTGCTGG	ATGTGGCTTG	ACTCTATGGA
15	3451	CTATGAGAAA	GCGCCACGCT	TCCCCAAGGG	AGAAAGGCGG
		GATACTCTTT	CGCGGTGCGA	AGGGCTTCCC	TCTTTCCGCC
	3501	GGTAAGCGGC	AGGGTCGGAA	CAGGAGAGCG	CACGAGGGAG
		CCATTTCGCG	TCCCAGCCTT	GTCTCTCTCG	GTGCTCCCTC
	3551	GAAACGCCTG	GTATCTTTAT	AGTCCTGTCT	GGTTTCGCCA
20		CTTTGCGGAC	CATAGAAATA	TCAGGACAGC	CCAAAGCGGT
	3601	GAGCGTCGAT	TTTTGTGATG	CTCGTCAGGG	GGGCGGAGCC
		CTCGCAGCTA	AAAACACTAC	GAGCAGTCCC	CCCGCCTCGG
	3651	CGCCAGCAAC	GCGGCCTTTT	TACGGTTCCCT	GGCCTTTTGC
		GCGGTTCGTT	CGCCGGAATA	ATGCCAAGGA	CCGGAAACG
25	3701	CTCACATGTT	CTTTCCTGCG	TTATCCCTTG	ATTCTGTGGA
		GAGTGTAACA	GAAAGGACGC	AATAGGGGAC	TAAGACACCT
	3751	ACCGCCTTTG	AGTGAGCTGA	TACCGCTCGC	CGCAGCCGAA
		TGGCGGAAAC	TCACTCGACT	ATGGCGAGCG	GCGTCGGCTT
	3801	CAGCGAGTCA	GTGAGCGAGG	AAGCGGAAGA	GCGCCCAATA
30		GTCTCGCTCAGT	CACTCGCTCC	TTTCGCTTCT	CGCGGGTTAT
	3851	GTCCTCCCGC	GCGTTGGCCG	ATTCAATTAAT	GCAGCTGGCA
		GAGAGGGGCG	CGCAACCGGC	TAAGTAATTA	CGTCGACCGT
	3901	CCCGACTGGA	AAGCGGGCAG	TGAGCGCAAC	GCAATTAATG
		GGGCTGACCT	TTTCGCCCCT	ACTCGCGTTG	CGTTAATTAC
35	3951	CACTCATTAG	GCACCCAGG	CTTTTACACTT	TATGCTTCCG
		GTGAGTAATC	CGTGGGGTCC	GAAATGTGAA	ATACGAAGGC
	4001	TGTGTGGAAT	TGTGAGCGGA	TAACAATTTT	ACACAGGAAA
		ACACACCTTA	ACACTCGCCT	ATTGTTAAAG	TGTGTCCTTT
	4051	CATGATTACG	AATTGAATTG	CGGCCGCAAT	TCTGAATGTT
40		GTACTAATGC	TTAACTTAAC	GCCGCGCTTA	AGACTTACAA
	4101	CTTTGGATGA	AGCTATAAAT	ATGCATTGGA	AAAATAATCC
		GAAACCTACT	TCGATATTTA	TACGTAACCT	TTTTATTAGG
	4151	AGGATTCAAA	TACTACAAAA	CCTAAGCGAT	AATATGTTAA
		TCCTAAGTTT	ATGATGTTTT	GGATTGCTTA	TTATACAATT
45	4201	TCTTAACGAC	GCTTTAAATA	TACACAAATA	AACATAATTT
		AGAATTGCTG	CGAAATTTAT	ATGTGTTTAT	TTGTATTAAA
	4251	TAACAAATAA	CTAAAACATA	AAAATAATAA	AAGGAAATGT
		ATTGTTTATT	GATTTTGTAT	TTTTATTATT	TTCTTTTACA
	4301	TTATTTTACT	CAGGAATGGG	GTTAAATATT	TATATCACGT
50		AATAAAATGA	GTCCTTACCC	CAATTTATAA	ATATAGTGCA
	4351	ACTGTTATCG	TATACTCTTT	ACAACTACTA	TTACGAATAT
		TGACAATAGC	ATATGAGAAA	TGTTAATGAT	AATGCTTATA
	4401	ATAAGATTAC	GTATTTAAGA	GAATCTTGTC	ATGATAATTG
		TATTCTAATG	CATAAATCTT	CTTAGAACAG	TACTATTAAC
55	4451	AGTGATAAAT	GCTATTTTCG	ATCGTTACAT	AAAGTCAGTT
		TCACTATTTA	CGATAAAGCG	TAGCAATGTA	TTTCAGTCAA
	4501	ATTTGACAGA	TGTAACCTTA	TAGGTGCAAA	AATGTTAAAT
		TAAACTGTCT	ACATTGAATT	ATCCACGTTT	TTACAATTTA
	4551	TATCGGAAGA	TAGGATACCA	GTTATATTAT	ACAAAAATCA

	ATAGCCTTCT	ATCCTATGGT	CAATATAATA	TGTTTTTAGT	GACCAACCTA
4601	AAAACAGATT	CTGCAATATT	CGTAAAAGAT	GAAGATTACT	GCGAATTTGT
	TTTTGTCTAA	GACGTTATAA	GCATTTTCTA	CTTCTAATGA	CGCTTAAACA
4651	AAACTATGAC	AATAAAAAGC	CATTTATCTC	AACGACATCG	TGTAATTTCTT
5	TTTGATACTG	TTATTTTTCG	GTAAATAGAG	TTGCTGTAGC	ACATTAAGAA
4701	CCATGTTTTA	TGTATGTGTT	TCAGATATTA	TGAGATTACT	ATAAACTTTT
	GGTACAAAAT	ACATACACAA	AGTCTATAAT	ACTCTAATGA	TATTTGAAAA
4751	TGTATACTTA	TATTCCTGTA	ACTATATTAA	TCATGAAGAA	AATGAAAAAG
	ACATATGAAT	ATAAGGCATT	TGATATAATT	AGTACTTCTT	TTACTTTTTC
10	4801	TATAGAAGCT	GTTTACGAGC	GGTTGTTGAA	AACAACAAAA
	ATATCTTCGA	CAAGTGCTCG	CCAACAACCT	TTGTTGTTTT	AATATGTAAG
4851	AAGATGGCTT	ACATATACGT	CTGTGAGGCT	ATCATGGATA	ATGACAATGC
	TTCTACCGAA	TGTATATGCA	GACACTCCGA	TAGTACCTAT	TACTGTACG
4901	ATCTCTAAAT	AGGTTTTCGG	ACAATGGATT	CGACCCTAAC	ACGGAATATG
15	TAGAGATTTA	TCCAAAAACC	TGTTACCTAA	GCTGGGATTG	TGCCTTATAC
4951	GTACTCTACA	ATCTCCTCTT	GAAATGGCTG	TAATGTTCAA	GAATACCGAG
	CATGAGATGT	TAGAGGAGAA	CTTTACCGAC	ATTACAAGTT	CTTATGGCTC
5001	GCTATAAAAA	TCTTGATGAG	GTATGGAGCT	AAACCTGTAG	TTACTGAATG
	CGATACTTTT	AGAACTACTC	CATACCTCGA	TTTGGACATC	AATGACTTAC
20	5051	CACAACCTCT	TGTCTGCATG	ATGCGGTGTT	GAGAGACGAC
	GTGTTGAAGA	ACAGACGTAC	TACGCCACAA	CTCTCTGCTG	ATGTTTTATC
5101	TGAAAGATCT	GTTGAAGAAT	AACTATGTAA	ACAATGTTCT	TTACAGCGGA
	ACTTTCTAGA	CAACTTCTTA	TTGATACATT	TGTTACAAGA	AATGTCGCCT
5151	GGCTTTTACT	CTTTGTGTTT	GGCAGCTTAC	CTTAACAAAG	TTAATTGCGT
25	CCGAAATGAG	GAAACACAAA	CCGTCGAATG	GAATTGTTTC	AATTAAACCA
5201	TAAACTTCTA	TTGGCTCATT	CGGCGGATGT	AGATATTTCA	AACACGGATC
	ATTTGAAGAT	AACCGAGTAA	GCCGCCTACA	TCTATAAAGT	TTGTGCCTAG
5251	GGTTAACTCC	TCTACATATA	GCCGTATCAA	ATAAAAAATT	AACATGGTTT
	CCAATTGAGG	AGATGTATAT	CGGCATAGTT	TATTTTTTAA	TTGTTACCAA
30	5301	AAACTTCTAT	TGAACAAAGG	TGCTGATACT	GACTTGCTGG
	TTTGAAGATA	ACTTGTTCCT	ACGACTATGA	CTGAACGACC	TATTTGACCC
5351	ATGTACTCCT	TTAATGATCG	CTGTACAATC	TGGAAATATT	GAAATATGTA
	TACATGAGGA	AATTACTAGC	GACATGTTAG	ACCTTTATAA	CTTTATACAT
5401	GCACACTACT	TAAAAAAAT	AAAATGTCCA	GAAGTGGGAA	AAATGTATCT
35	CGTGTGATGA	ATTTTTTTTA	TTTTACAGGT	CTTGACCCTT	TTTAACTAGA
5451	TGCCAGCTGT	AATTCATGGT	AGAAAAGAAG	TGCTCAGGCT	ACTTTTCAAC
	ACGGTCGACA	TTAAGTACCA	TCTTTTCTTC	ACGAGTCCGA	TGAAAAGTTG
5501	AAAGGAGCAG	ATGTAAACTA	CATCTTTGAA	AGAAATGGAA	AATCATATAC
	TTTCCTCGTC	TACATTTGAT	GTAGAAACTT	TCTTTACCTT	TTAGTATATG
40	5551	TGTTTTGGAA	TTGATTAAAG	AAAAGTTACTC	TGAGACACAA
	ACAAAACCTT	AACTAATFTC	TTTCAATGAG	ACTCTGTGTT	TTCTCCATCG
5601	TGAAGTGGTA	CTCTCAAAGG	TACGTGACTA	ATTAGCTATA	AAAAGGATCC
	ACTTCACCAT	GAGAGTTTCC	ATGCACTGAT	TAATCGATAT	TTTTCCCTAGG
5651	TAGAGGATCA	TTATTTAACG	TAAACTAAAT	GGAAAAGCTA	TTTACAGGTA
45	ATCTCCTAGT	AATAAATTGC	ATTTGATTTA	CCTTTTCGAT	AAATGTCCAT
5701	CATACGGTGT	TTTCTGGAAT	CAAAATGATTC	TGATTTTGAG	GATTTTATCA
	GTATGCCACA	AAAGACCTTA	GTTTACTAAG	ACTAAAACCTC	CTAAAAAGT
5751	ATACAATAAT	GACAGTGCTA	ACTGGTAAAA	AAGAAAGCAA	ACAATTATCA
	TATGTTATTA	CTGTCACGAT	TGACCATTTT	TTCTTTTCGTT	TGTTAATAGT
50	5801	TGGCTAACAA	TTTTTATTAT	ATTTGTAGTA	TGCATAGTGG
	ACCGATTGTT	AAAAATAATA	TAAACATCAT	ACGTATCACC	AGAAATGCAA
5851	TCTTTATTTA	AAGTTAATGT	GTTAAGATTA	AATGGAGTAA	TTGGATCCCC
	AGAAATAAAT	TTCAATTACA	CAATTCTAAT	TTACCTCATT	AACCTAGGGG
5901	CATCGATGGG	GAATTCACCTG	GCCGTGCTTT	TACAACGTCG	TGACTGGGAA
55	GTAGCTACCC	CTTAAGTGAC	CGGCAGCAAA	ATGTTGCAGC	ACTGACCCTT
5951	AACCTGGCG	TTACCCAACCT	TAATCGCCTT	GCAGCACATC	CCCCTTTCGC
	TTGGGACCGC	AATGGGTTGA	ATTAGCGGAA	CGTCGTGTAG	GGGGAAAGCG
6001	CAGCTGGCGT	AATAGCGAAG	AGGCCCGCAC	CGATCGCCCT	TCCCAACAGT
	GTGACCGCA	TTATCGCTTC	TCCGGGCGTG	GCTAGCGGGA	AGGGTTGTCA

6051	TGCGCAGCCT	GAATGGCGAA	TGGCGCTTTG	CCTGGTTTCC	GGCACCAGAA
	ACGCGTCGGA	CTTACCGCTT	ACCGCGAAAC	GGACCAAAGG	CCGTGGTCTT
6101	GCGGTGCCGG	AAAGCTGGCT	GGAGTGCAT	CTTCCTGAGG	CCGATACTGT
	CGCCACGGCC	TTTCGACCGA	CCTCACGCTA	GAAGGACTCC	GGCTATGACA
5	6151	CGTCGTCCCC	TCAAAC TGCG	AGATGCACGG	TTACGATGCG
		GCAGCAGGGG	AGTTTGACCG	TCTACGTGCC	AATGCTACGC
	6201	CCAACGTAAC	CTATCCCAT	ACGGTCAATC	CGCCGTTTGT
		GGTTGCATTG	GATAGGGTAA	TGCCAGTTAG	GCGGCAAACA
	6251	AATCCGACGG	GTTGTTACTC	GCTCACATTT	AATGTTGATG
10		TTAGGCTGCC	CAACAATGAG	CGAGTGTAAC	TTTCACTAC
	6301	ACAGGAAGGC	CAGACGCGAA	TTATTTTTGA	TGGCGTTAAC
		TGTCCTTCCG	GTCTGCGCTT	AATAAAAACT	ACCGCAATTG
	6351	ATCTGTGGTG	CAACGGGCGC	TGGGTGCGTT	ACGGCCAGGA
		TAGACACCAC	GTTGCCCGCG	ACCCAGCCAA	TGCCGGTCCT
15	6401	CCGTCTGAAT	TTGACCTGAG	CGCATTTTTA	CGCGCCGGAG
		GGCAGACTTA	AACCTGGACTC	GCGTAAAAAT	GCGCGGCCTC
	6451	CGCGGTGATG	GTGCTGCGTT	GGAGTGACGG	CAGTTATCTG
		GCGCCACTAC	CACGACGCAA	CCTCACTGCC	GTCAATAGAC
	6501	ATATGTGGCG	GATGAGCGCG	ATTTTCCGTG	ACGTCTCGTT
20		TATACACCGC	CTACTCGCCG	TAAAAGGCAC	TGCAGAGCAA
	6551	CCGACTACAC	AAATCAGCGA	TTTCCATGTT	GCCACTCGCT
		GGCTGATGTG	TTTAGTTCGT	AAAGGTACAA	CGGTGAGCGA
	6601	TTTCAGCCGC	GCTGTACTGG	AGGCTGAAGT	TCAGATGTGC
		AAAGTCGGCG	CGACATGACC	TCCGACTTCA	AGTCTACACG
25	6651	GTGACTACCT	ACGGGTAAAC	GTTCCTTTAT	GGCAGGGTGA
		CACCTGATGA	TGCCCCATTG	CAAAAGAAATA	CCGTCCCACT
	6701	GCCAGCGGCA	CCGCGCCTTT	CGGCGGTGAA	ATTATCGATG
		CGGTGCGCGT	GGCGCGGAAA	GCCGCCACTT	TAATAGCTAC
	6751	TTATGCCGAT	CGCGTCACAC	TACGTCTGAA	CGTCGAAAAC
30		AATACGGCTA	GCGCAGTG TG	ATGCAGACTT	GCAGCTTTTG
	6801	GGAGCGCCGA	AATCCCGAAT	CTCTATCGTG	CGGTGCGTTGA
		CCTCGCGGCT	TTAGGGCTTA	GAGATAGCAC	GCCACCAACT
	6851	GCCGACGGCA	CGCTGATTGA	AGCAGAAGCC	TGCGATGTGC
		CGGCTGCCGT	GCGACTAACT	TCGTCTTCGG	ACGCTACAGC
35	6901	GGTGCGGATT	GAAAATGGTC	TGCTGTCTGCT	GAACGGCAAG
		CCACGCCTAA	CTTTTACCAG	ACGACGACGA	CTTGCCGTTT
	6951	TTGAGGCGCT	TAACCGTCAC	GAGCATCATC	CTCTGCATGG
		AAGCTCCGCA	ATTGGCAGTG	CTCGTAGTAG	GAGACGTACC
	7001	GATGAGCAGA	CGATGGTGCA	GGATATCCTG	CTGATGAAGC
40		CTACTCGTCT	GCTACCACGT	CCTATAGGAC	GACTACTTCG
	7051	TAACGCCCGTG	CGCTGTTTCG	ATTATCCGAA	CCATCCGCTG
		ATTGCGGCAC	GCGACAAGCG	TAATAGGCTT	GGTAGGCGAC
	7101	TGTGCGACCG	CTACGGCCTG	TATGTGGTGG	ATGAAGCCAA
		ACACGCTGGC	GATGCCGGAC	ATACACCACC	TACTTCGGTT
45	7151	CACGGCATGG	TGCCAATGAA	TCGTCTGACC	GATGATCCGC
		GTGCCGTACC	ACGGTTACTT	AGCAGACTGG	CTACTAGGCG
	7201	GGCGATGAGC	GAACGCGTAA	CGCGAATGGT	GCAGCGCGAT
		CCGCTACTCG	CTTGCGCAT	GCGCTTACCA	CGTCGCGCTA
	7251	CGAGTGTGAT	CATCTGGTCTG	CTGGGGAAATG	AATCAGGCCA
50		GCTCACACTA	GTAGACCAGC	GACCCCTTAC	TTAGTCCGGT
	7301	CACGACGCGC	TGTATCGCTG	GATCAAATCT	GTCGATCCTT
		GTGCTGCGCG	ACATAGCGAC	CTAGTTTAGA	CAGCTAGGAA
	7351	GCAGTATGAA	GGCGGCGGAG	CCGACACCAC	GGCCACCGAT
		CGTCATACTT	CCGCCGCCCTC	GGCTGTGGTG	CCGGTGGCTA
55	7401	CGATGTACGC	GCGCGTGGAT	GAAGACCAGC	CCTTCCCGGC
		GCTACATGCG	CGCGCACCTA	CTTCTGGTCTG	GGAAGGGCCG
	7451	TGGTCCATCA	AAAAATGGCT	TTCGCTACCT	GGAGAGACGC
		ACCAGGTAGT	TTTTTACCGA	AAGCGATGGA	CCTCTCTGCG
	7501	CCTTTGCGAA	TACGCCACCG	CGATGGGTAA	CAGTCTTGGC

		GGAACGCTT	ATGCGGGTGC	GCTACCCAT	GTCAGAACCG	CCAAAGCGAT
7551		AATACTGGCA	GGCGTTTCGT	CAGTATCCCC	GTTTACAGGG	CGGCTTCGTC
		TTATGACCGT	CCGCAAAGCA	GTCATAGGGG	CAAATGTCCC	GCCGAAGCAG
7601		TGGGACTGGG	TGGATCAGTC	GCTGATTAAA	TATGATGAAA	ACGGCAAACCC
5		ACCTGACCC	ACCTAGTCAG	CGACTAATTT	ATACTACTTT	TGCCGTTGGG
		GTGGTCGGCT	TACGGCGGTG	ATTTTGGCGA	TACGCCGAAC	GATCGCCAGT
7651		CACCAGCCGA	ATGCCGCCAC	TAAAACCGCT	ATGCGGCTTG	CTAGCGGTCA
		TCTGTATGAA	CGGTCTGGTC	TTTGCCGACC	GCACGCCGCA	TCCAGCGCTG
7701		AGACATACTT	GCCAGACCAG	AAACGGCTGG	CGTGCGGCGT	AGGTCGCGAC
10		7751	ACGGAAGCAA	AACACCAGCA	GCAGTTTTTC	CAGTTCGGTT
			TGCCTTCGTT	TTGTGGTTCGT	CGTCAAAAAG	GTCAAGGCAA
		7801	AACCATCGAA	GTGACCAGCG	AATACCTGTT	CCGTTCATAGC
			TTGGTAGCTT	CACCTGGTCGC	TTATGGACAA	GGCAGTATCG
		7851	TCTGCACCTG	GATGGTGGCG	CTGGATGGTA	AGCCGCTGGC
15			AGGACGTGAC	CTACCACCGC	GACCTACCAT	TCCGCGACCG
		7901	GTGCCTCTGG	ATGTCGCTCC	ACAAGGTAAA	CAGTTGATTG
			CACCGAGACC	TACAGCGAGG	TGTTCCATTT	GTCAACTAAC
		7951	ACTACCGCAG	CCGGAGAGCG	CCGGGCAACT	CTGGCTCACA
			TGATGGCGTC	GGCCTCTCGC	GGCCCGTTGA	GACCGAGTGT
20		8001	TGCAACCGAA	CGCGACCGCA	TGGTCAGAAG	CCGGGCACAT
			ACGTTGGCTT	GCGCTGGCGT	ACCAGTCTTC	GGCCCGTGTA
		8051	CAGCAGTGGC	GTCTGGCGGA	AAACCTCAGT	GTGACGCTCC
			GTCCTCACCG	CAGACCGCCT	TTTGGAGTCA	CACTGCGAGG
		8101	CCACGCCATC	CCGCATCTGA	CCACCAGCGA	AATGGATTTT
25			GGTGCGGTAG	GGCGTAGACT	GGTGGTCGCT	TTACCTAAAA
		8151	TGGGTAATAA	GCGTTGGCAA	TTTAACCGCC	AGTCAGGCTT
			ACCCATTATT	CGCAACCGTT	AAATTGGCGG	TCAGTCCGAA
		8201	ATGTGGATTG	GCGATAAAAA	ACAACCTGCTG	ACGCCGCTGC
			TACACCTAAC	CGCTATTTTT	TGTTGACGAC	TGCGGCGACG
30		8251	CACCCGTGCA	CCGCTGGATA	ACGACATTGG	CGTAAGTGAA
			GTGGGCACGT	GGCGACCTAT	TGCTGTAACC	GCATTCACTT
		8301	TTGACCCCTAA	CGCCTGGGTC	GAACGCTGGA	AGGCGGCGGG
			AACTGGGATT	GCGGACCCAG	CTTGCGACCT	TCCGCCGCCC
		8351	GCCGAAGCAG	CGTTGTTGCA	GTGCACGGCA	GATACACTTG
35			CGGCTTCGTC	GCAACAACGT	CACGTGCCGT	CTATGTGAAC
		8401	GCTGATTACG	ACCGCTCACG	CGTGGCAGCA	TCAGGGGAAA
			CGACTAATGC	TGGCGAGTGC	GCACCGTCGT	AGTCCCCCTT
		8451	TCAGCCGGAA	AACTTACCGG	ATTGATGGTA	GTGGTCAAAT
			AGTCGGCCTT	TTGGATGGCC	TAACCTACCAT	CACCAGTTTA
40		8501	GTTGATGTTG	AAGTGGCGAG	CGATACACCG	CATCCGGCGC
			CAACTACAAC	TTACCCGCTC	GCTATGTGGC	GTAGGCCGCG
		8551	GAACCTGCCAG	CTGGCGCAGG	TAGCAGAGCG	GGTAACTGG
			CTTGACGGTC	GACCGCGTCC	ATCGTCTCGC	CCATTTGACC
		8601	GGCCGCAAGA	AAACTATCCC	GACCGCCTTA	CTGCCGCCTG
45			CCGGCGTCTT	TTTGATAGGG	CTGGCGGAAT	GACGGCGGAC
		8651	TGGGATCTGC	CATTGTCAGA	CATGTATACC	CCGTACGTCT
			ACCTTAGACG	GTAACAGTCT	GTACATATGG	GGCATGCAGA
		8701	AAACGGTCTG	CGCTGCGGGA	CGCGCGAATT	GAATTATGGC
			TTTGCCAGAC	GCGACGCCCT	GCGCGCTTAA	CTTAATACCG
50		8751	GGCGCGGCGA	CTTCCAGTTC	AACATCAGCC	GCTACAGTCA
			CCGCGCCGCT	GAAGGTCAAG	TTGTAGTCGG	CGATGTCAGT
		8801	ATGGAAACCA	GCCATCGCCA	TCTGCTGCAC	GCGGAAGAAG
			TACCTTTGGT	CGGTAGCGGT	AGACGACGTG	CGCCTTCTTC
		8851	GAATATCGAC	GGTTTCCATA	TGGGGATTGG	TGGCGACGAC
55			CTTATAGCTG	CCAAAGGTAT	ACCCCTAACC	ACCGCTGCTG
		8901	CGTCAGTATC	GGCGGAATTC	CAGCTGAGCG	CCGGTCGCTA
			GCAGTCATAG	CCGCCTTAAG	GTCGACTCGC	GGCCAGCGAT
		8951	TTGGTCTGGT	GTCAAAAATA	ATAATAACCG	GGCAGGGGGG
			AACCAGACCA	CAGTTTTTAT	TATTATTGGC	CCGTCCCCCC
						TAGGCCTCGA

9001	TATCGCAGAT	CAATGATCGC	TGTACAATCT	GGAAATATTG	AAATATGTAG
	ATAGCGTCTA	GTTACTAGCG	ACATGTTAGA	CCTTTATAAC	TTTATACATC
9051	CACACTACTT	AAAAAAAATA	AAATGTCCAG	AACTGGGAAA	AATTGATCTT
	GTGTGATGAA	TTTTTTTTTAT	TTTACAGGTC	TTGACCCTTT	TTAACTAGAA
5	9101	GCCAGCTGTA	ATTCATGGTA	GAAAAGAAGT	GCTCAGGCTA
		CGGTTCGACAT	TAAGTACCAT	CTTTTCTTCA	CGAGTCCGAT
	9151	AAGGAGCAGA	TGTAAACTAC	ATCTTTGAAA	GAAATGGAAA
		TTCCCTCGTCT	ACATTTGATG	TAGAAACTTT	CTTTACCTTT
	9201	GTTTTGGAAT	TGATTAAAGA	AAGTTACTCT	GAGACACAAA
10		CAAAACCTTA	ACTAATTTCT	TTCAATGAGA	CTCTGTGTTT
	9251	GAAGTGGTAC	TCTCAAAGGT	ACGTGACTAA	TTAGCTATAA
		CTTCACCATG	AGAGTTTCCA	TGCACTGATT	AATCGATATT
	9301	GTACCCTCGA	GTCTAGAATC	GATCCCGGGT	TAATTAATTA
		CATGGGAGCT	CAGATCTTAG	CTAGGGCCCA	ATTAATTAAT
15	9351	AAGGTGAAAA	CGAAACTATT	TGTAGCTTAA	TTAATTAGAG
		TTCCACTTTT	GCTTTGATAA	ACATCGAATT	AATTAACTCT
	9401	CTATACTTAA	AAAGTGAAAA	TAAATACAAA	GGTTCTTGAG
		GATATGAATT	TTTCACTTTT	ATTTATGTTT	CCAAGAACTC
	9451	AATTGAAAGC	GAGAAATAAT	CATAAATTAT	TTCAATTATCG
20		TTAACTTTTCG	CTCTTTATTA	GTATTTAATA	AAGTAATAGC
	9501	TAAGTTTGTA	TCGTA		GCTATAGGCA
		ATTCAAACAT	AGCAT		

**FIGURE 6**